Turnover of strain-level diversity modulates functional traits in the honeybee gut microbiome between nurses and foragers

Gilles LC Baud, Kirsten M Ellegaard, Philipp Engel

Department of Fundamental Microbiology, University of Lausanne, 1015 Lausanne, Switzerland

Correspondence:

Prof. Philipp Engel

Department of Fundamental Microbiology

University of Lausanne, CH-1015 Lausanne, Switzerland

Tel.: +41 (0)21 692 56 12

e-mail: philipp.engel@unil.ch

Abstract

Strain-level diversity is widespread among bacterial species and can expand the functional potential of natural microbial communities. However, to what extent communities shift in strain composition in response to environmental changes has mostly remained elusive. Here, we used shotgun metagenomics to compare the gut microbiota of two behavioral states of Western honeybees, larvae-feeding nurses and pollen foragers. While their gut microbiota is composed of the same bacterial species, we detected consistent changes in strain-level composition between nurses and foragers. Single nucleotide variant profiles of predominant bacterial species clustered by behavioral state rather than by colony or sampling location. Moreover, we identified strain-specific gene content related to nutrient utilization, vitamin biosynthesis, and cell-cell interactions specifically associated with the two behavioral states. Our findings show that strain-level diversity in host-associated communities can turnover in response to environmental changes modulating functional traits of communities and highlighting that strains are ecologically relevant units.

Introduction

Microbial communities are highly dynamic. They undergo compositional changes in response to environmental factors, which can profoundly influence their functions. For example, antibiotic treatment can disturb the human gut microbiota, altering its effect on the host metabolism and immune system, and facilitating the invasion of pathogens (Keeney et al., 2014). Such microbiota shifts are difficult to predict as they often occur irregularly and in different contexts. However, when linked to recurrent or predetermined environmental changes, microbial communities can undergo consistent compositional transitions, providing excellent opportunities to study ecological interactions governing community shifts. For example, bacterioplankton communities in the ocean that depend on sunlight intensity (Martin-Platero et al., 2018; Teeling et al., 2016, 2012), or gut microbial communities of animals that have a seasonal diet, or hibernate during the winter months (Davenport et al., 2014; Maurice et al., 2015; Smits et al., 2017; Sommer et al., 2016), turnover as a function of season. Similarly, the assembly of the human gut microbiota after birth follows a predictable pattern characterized by the succession of different microbial taxa during the first few months of life (Bäckhed et al., 2015; Rao et al., 2021).

Compositional changes associated with the turnover of microbial communities have mostly been studied at the genus or species level using 16S rRNA gene analysis. However, natural microbial communities usually harbor a large extent of strain-level diversity, with strains of the same species carrying distinct functions. In the human gut, strain-level differences can influence the metabolism of dietary compounds (Barboza et al., 2009; Patnode et al., 2021) or drugs (Haiser et al., 2013), or modulate their interactions with the host immune system (Bron et al., 2012). Moreover, several recent studies have suggested that strains of the same

55 species can exhibit different dynamics and govern the eco-evolutionary dynamics in microbial
56 communities (Goyal et al., 2022; Leventhal et al., 2018). However, if natural microbial
57 communities undergo consistent compositional transitions at the strain-level in response to
58 environmental changes is less well understood.

59 Honeybees (*Apis mellifera*) can be useful models for studying strain-level dynamics in host-
60 associated microbial communities. They are eusocial animals that live in large colonies
61 composed of a single queen, a few hundred male bees, and thousands of sterile female
62 worker bees. They harbor a relatively simple, but highly specialized microbial community in
63 their gut, composed of <10 bacterial phylotypes (i.e. bacteria sharing >97% sequence identity
64 in the 16S rRNA gene)(Bonilla-Rosso and Engel, 2018; Kwong and Moran, 2016). Most of these
65 phylotypes have diverged into sequence discrete populations (sequence clusters of >95% ANI;
66 hereafter referred to as species), each harboring a high degree of strain-level diversity and
67 variation in functional gene content (Ellegaard et al., 2019, 2019; Ellegaard and Engel, 2019;
68 Engel et al., 2014; Segers et al., 2017; Zheng et al., 2019, 2016).

69 Using shotgun metagenomics, we have previously shown that while phylotypes, and species
70 within phylotypes, co-exist in individual bees, strains within species segregate, suggesting that
71 each worker bee harbors a distinct community at the strain-level (Ellegaard and Engel, 2019).
72 These differences in strain-level composition could not be attributed to any variation in age,
73 time of sampling, or colony affiliation in our previous study. However, only a few colonies
74 were sampled, and bees were exclusively sampled from within the hive between September
75 and February raising the possibility that most microbiota samples originated from nurse bees
76 or long-lived winter bees. During the summer months, worker bees undergo an age-related
77 behavioral maturation (Brodschneider and Crailsheim, 2010; Seeley, 1985; Smith et al., 2008).

Young nurse bees feed on pollen and fulfill tasks inside the hive related to brood care, nest building, and food processing. However, when 2-3 weeks old nurses transition to forager bees, which feed on honey, and collect pollen and nectar outside the hive. Besides the dramatic change in dietary intake, this behavioral maturation is accompanied by a series of physiological changes, such as endocrine secretion, fat body metabolism, and brain activity (Robinson, 2002; Whitfield et al., 2006). Community profiling based on 16S rRNA gene sequencing has shown that the nurse-to-forager transition is associated with a consistent change in gut microbiota composition (Jones et al., 2018; Kešnerová et al., 2020). While the microbiota of both bee types is dominated by the same bacterial phylotypes, the composition of the community shifts with a marked decrease in the absolute abundance of most phylotypes in forager versus nurse bees.

Here, we tested if the gut microbiota of honeybees during the transition from nurse to forager bees also undergoes consistent compositional changes at the strain-level. To this end, we analyzed shotgun metagenomics data from nurse and forager bees sampled across 15 colonies in Western Switzerland. Our results show that the two behavioral states harbor distinctive bacterial communities across colonies, characterized by consistent shifts in strain composition and functional gene content of major bacterial species. These findings show that strains are ecologically distinct units that undergo a predictable ecological transition when bees turn from nurses to foragers. We speculate that this microbiota shift is driven by nutrient availability disparities or physiological changes in the gut of foragers relative to nurses.

99 ## Results

100 **Bacterial loads and species-level community composition differ between nurse and**
101 **forager bees**

102 To investigate compositional changes with genomic resolution in the honeybee gut
103 microbiota in response to behavioral state, we collected metagenomic samples from larvae-
104 feeding nurses and pollen foragers (**Supplementary Table 1**). In contrast to previous studies
105 (Ellegaard et al., 2020; Ellegaard and Engel, 2019; Engel et al., 2012; Wu et al., 2021; Zheng et
106 al., 2019), samples consisted of pooled honeybee guts to ensure representative sampling of
107 the strain-level diversity for each colony. In total, we sampled 15 colonies from five different
108 sampling locations (three each), which allowed us to account for eventual variation between
109 sites and colonies, not related to behavioral state.

110 Due to the different dietary habits of the two behavioral states, the hindguts of nurses were
111 usually pollen-filled, as indicated by the enlarged gut size, and the opacity and yellow color of
112 the gut content. In contrast, the hindguts of foragers were smaller and contained mostly
113 nectar based on their translucent appearance. Accordingly, nurse guts were on average 2.9x
114 heavier than their forager counterparts (p=3.3e-09, Wilcoxon sign rank test, **Supplementary**
115 **Figure 1)**.

116 To estimate the absolute abundance of bacteria relative to the amount of host tissue in each
117 pooled gut sample, we performed qPCR on genomic DNA targeting the bacterial 16S rRNA
118 gene. This revealed that the bacterial loads in the forager samples were on average 2.9x lower
119 than the ones in the nurse samples (**Figure 1A**, p=7.1e-05, Wilcoxon sign rank test), confirming
120 previous results by Kešnerová et al. (2020).

121 Shotgun metagenomics sequencing resulted in 58.4 – 76.7 million reads per sample. To
122 compare the bacterial community composition of nurses and foragers, the metagenomic
123 reads were mapped to a reference database of isolate genomes obtained from the gut
124 microbiota of social bees, as in Ellegaard et al. (2020) and Ellegaard and Engel (2019). The
125 non-redundant database was composed of 212 isolate genomes and encompassed all major
126 community members described to be present in the honeybee gut microbiota. Between 19.8
127 and 57.4 % of all metagenomic reads mapped to the bacterial reference database. Of the
128 remaining reads, 30.6 - 76.0 % mapped to the host genome of *A. mellifera,* and only 4.2 - 20.8
129 % of the reads could not be assigned (**Supplementary Figure 2**). Forager samples overall had
130 a lower proportion of reads mapping to the bacterial database than the nurse samples
131 (forager average 35.6%, nurse average 50.8%, p=6.1e-05, Wilcoxon sign rank test), which is in
132 line with the difference in bacterial loads between the two behavioral states.

133 We next determined the relative and absolute abundance of the major phylotypes and
134 species in each sample based on the read coverage of single-copy core genes and the total
135 number of 16 rRNA gene copies determined by qPCR. We will refer to different species within
136 phylotypes using the previously established SDP numbering (Ellegaard et al., 2020; Ellegaard
137 and Engel, 2019), for reasons of readability but also because only for a subset of them proper
138 species names have been established to date. A list with all SDPs and corresponding species
139 names can be found in **Supplementary Table 2**. Both the phylotype and species-level
140 composition were comparable to the community profiles detected in our previous shotgun
141 metagenomic analyses which were based on individual bees (Ellegaard et al., 2020; Ellegaard
142 and Engel, 2019). Firmicutes (Bombilactobacilli and Lactobacilli) and Bifidobacteria
143 dominated the community, and for the phylotypes Bombilactobacillus, Lactobacillus,
144 Bifidobacterium, and Gilliamella, different species were detected in all samples (**Figure 1B,**

145 **Supplementary Figure 3**). Consistent with previous findings that social bees harbor host-
146 specific communities (Kwong et al., 2017; Kwong and Moran, 2016; Moran et al., 2019), most
147 reads mapped to species previously detected in *A. mellifera*. Only a few reads (0.15% - 3.1%)
148 in nine samples also mapped to bacteria from other bee species (including Lactobacilli
149 isolated from *Apis cerana,* and Gilliamella isolated from bumble bees and *Apis cerana*, **see**
150 **Supplementary Table 3**)*,* which may also be due to spurious mapping*.*

151 Non-metric multidimensional scaling (NMDS) using Bray-Curtis dissimilarity of the bacterial
152 relative abundances separated nurse from forager samples at both the phylotype and the
153 species level (**Figure 1C and 1D**). In terms of absolute abundances (as determined by
154 multiplying the proportions of each community member by the total bacterial biomass in each
155 sample, see methods), most community members were less abundant in foragers than in
156 nurse bees (**Supplementary Figure 4**), a direct consequence of the lower total bacterial load
157 detected in forager samples. However, there were clear differences in the strength of these
158 changes between community members, even among species of the same phylotype (**Figure**
159 **1E**). For example, the total abundances of Lactobacilli Firm5_1 and Firm5_4 changed to a
160 lesser extent (average log2 fold change = -1.45 and -0.71) than those of Firm5_2 and Firm5_3
161 (average log2 fold change = -2.82 and -2.99) in foragers relative to nurses. This explains the
162 consistent shift towards a less even species composition of this phylotype in foragers
163 compared to nurses (**Figure 1B, Supplementary Figure 4**). Likewise, Bifido_1.2 was more
164 affected (average log2 fold change = -2.32) than any of the other species of the
165 Bifidobacterium phylotype (average log2 fold change across species of the Bifidobacterium
166 phylotype = -0.37) by the behavioral state transition of the host. Notably, four community
167 members (Bapis, Snod_1, Com_1 and Lkun) did not change in absolute bacterial load between

168 nurses and forager samples, suggesting that they are not affected in their abundance by the
169 behavioral transition.

170 In summary, although the gut microbiota of forager and nurse bees were composed of the
171 same community members, we found consistent differences in the absolute abundance of
172 many of them, not only at the phylotype level as previously shown (Jones et al., 2018;
173 Kešnerová et al., 2020), but also at the species level. Moreover, some species of a given
174 phylotype were more strongly affected than others suggesting species-specific responses to
175 the nurse-to-forager transition.

176

177 **Consistent differences in strain-level composition between nurse and forager bees**

178 Previous studies have shown that most honeybee gut microbiota species harbor considerable
179 levels of strain-level diversity (Ellegaard et al., 2020; Ellegaard and Engel, 2019). To test
180 whether some strains were preferentially associated with nurses or foragers, we mapped the
181 metagenomic reads against a reduced reference database containing one genome per species
182 and quantified the number of polymorphic sites per sample on the core genes as a measure
183 for the extent of strain-level diversity.

184 We detected between 1.53 and 8.53% polymorphic sites per species per sample, with
185 Firm5_4 showing the highest, and Firm4_2 the lowest average proportion of polymorphic
186 sites per sample (**Figure 2A**). These differences in proportions of polymorphic sites among
187 community members are comparable to those previously found by Ellegaard et al. (2020) and
188 Ellegaard and Engel (2019). To assess the total amount of strain-level diversity detected per
189 species in our study, we calculated cumulative proportions of polymorphic sites as a function
190 of the number of bee samples. The cumulative proportions of polymorphic sites across all

191 samples ranged between 4.35 and 11.76% per species, with Bifido_1.1 showing the highest
192 and Firm4_2 the lowest level of total diversity (**Figure 2B**). In all cases, we observed that the
193 slope of the curves rapidly decreased with the addition of new samples, indicating that many
194 of the polymorphic sites are shared across samples.

195 Four of the 17 analyzed species showed higher levels of strain-level diversity in foragers than
196 in nurse bees (**Figure 2A**): Bifido_1.2 (q-value=1.35e-4), Firm5_3 (q-value=2.15e-2), Gilli_2 (q-
197 value=1.35e-4), Snod_1 (q-value=8.36e-5). In contrast none of the species harbored higher
198 strain-level diversity in nurses than foragers indicating that diversity within bees tends to
199 increase after the transition. The forager samples had on average 2.9 times lower bacterial
200 load than nurse samples, but on average only 1.5 times smaller number of reads per sample.
201 This means that a larger fraction of the total bacterial population was analyzed for forager
202 relative to nurse samples, a difference that could potentially affect diversity estimates.
203 However, no significant correlation was found between the proportion of polymorphic sites
204 detected and the read coverage of our samples (**Supplementary Figure 5**), suggesting that we
205 had sufficiently deep coverage across the study.

206 Next, we assessed the difference in strain profiles between foragers and nurse bee samples
207 by comparing the distribution of polymorphic sites detected for each species across samples.
208 To do so, we computed the Jaccard distances between all sample pairs based on shared
209 polymorphic sites. Principal coordinate analysis of these similarity indices resulted in
210 clustering of the samples by host behavioral state for seven of the 17 analyzed species
211 (Permanova, q-value < 0.05, **Figure 2C**), including all species of Bifido_1 except Bifido_1.5,
212 Firm5_3, Firm5_4, Gilli_1, Gilli_2, and Snod_1. Notably, we also observed clustering by
213 sampling site in the principal coordinate analysis (Permanova, q-value < 0.05, **Supplementary**

214 **Figure 6**) for almost all species, with the exception of Bifido_1.3 and Bifido_1.5. However, for
215 the species that clustered by both factors (sampling site and behavioral state) the pseudo F-
216 statistic was generally larger for the behavioral state, suggesting a stronger influence than the
217 sampling site on the strain composition of the gut microbiota.

218 In summary, while previous publications could find no influence of worker bee age
219 (Ellegaard and Engel, 2019), we found consistent changes in terms of strain composition for
220 several species depending on the behavioral state of honeybee workers.

221

222 **Differences in gene content composition between nurse and forager bees**

223 Because nurses and foragers have a different physiology and diet, we hypothesized that
224 functional differences among strains of the same species may drive the consistent changes in
225 strain-level composition. Previous studies have shown that most species of the honeybee gut
226 microbiota harbor large accessory gene pools and that individual bees vary in terms of the
227 functional gene content encoded in a given species (Ellegaard and Engel, 2019).

228 To identify strain-specific genes that may be enriched in one behavioral state or the other,
229 we compared the read coverage of all genes assigned to a given species. Strain-specific genes
230 should have a lower coverage than the core genes of a given species across samples; and
231 those that predominate in one behavioral state relative to the other should show a consistent
232 coverage discrepancy between forager and nurse samples.

233 To implement this approach, we assembled all reads not mapping to the host genome and
234 binned the resulting contigs into bacterial species bins using our reference database (see
235 methods for details and **Supplementary Figure 7**). Then, we clustered the ORFs encoded in

236 the binned contigs into species-specific ortholog groups (OGs) and identified those OGs that
237 showed consistent coverage discrepancy between forager and nurse samples. In total, 81.7%
238 (65.3 – 90.9% per sample) of all assembled bases could be binned and 83.4% (73.2 – 86.1%
239 per sample) of the non-host reads mapped to the binned contigs indicating that we had
240 included most of the bacterial diversity present in the samples. Out of 45,863 species-specific
241 OGs, 1,551 revealed significant differences between the nurse and forager samples (q-value
242 < 0.05, Wilcoxon sign rank test, and absolute log2 fold change > 1, **Supplementary Dataset**
243 **1**), which indicated that these genes were more abundant in one host phenotype versus the
244 other. Notably, the number of differential abundant OGs substantially varied between species
245 ranging from 15 to 231 (**Figure 3A, Supplementary Dataset 1**). For example, in Bifido_1.1, we
246 detected very few consistent gene content differences between foragers and nurses (6 and 9
247 OGs, respectively) even though this species had the highest proportion of polymorphic sites
248 with a strong clustering of the samples by behavioral state. The largest differences in gene
249 content between nurses and foragers were found for three Lactobacillus Firm5 species
250 (Firm5_2: 184 OGs, Firm5_3: 131 OGs, Firm5_4: 177 OGs), two Gilliamella species (Gilli_1: 64
251 OGs, Gilli_2: 161 OGs), Snod_1 (197 OGs), and Bifido_1.2 (159 OGs). Consistent with these
252 findings, all of these species, except for Firm5_2, also showed clustering by behavioral state
253 in terms of polymorphic sites (**Figure 2C**). Interestingly, a larger fraction of the state-specific
254 OGs were found in foragers than in nurses in total (1019 versus 445 significantly different
255 OGs, respectively), but also for most of the individual species.

256 To determine the functions encoded by the behavioral state-specific gene content, we
257 annotated all OGs using eggnog and used COG categories as a first level of functional
258 characterization. Of all behavioral state-specific OGs, 33.1% had no COG annotation, while
259 17.3% of them fell into the COG category S (unknown function). For the rest of the significant

260 OGs, we categorized 20.6% of them as metabolism-related, 11.7% as linked to cellular
261 processes, signaling and structure, and 18.3% as information storage and processing **(Figure**
262 **3B**). Among the significant categorized OGs (excluding COG S), category G (carbohydrate
263 metabolism and transport) was the most important category, representing 11.4% of all
264 significant OGs. However, the proportion of different COG categories contributing to the
265 significant OGs varied between species **(Figure 3B**). In case of Firm4_1, COG category H
266 (coenzyme metabolism) was specifically enriched in the behavioral state-specific OGs relative
267 to the total number of OGs per species (Fisher's Exact test q-value=2.5e-06, **Supplementary**
268 **Table 4.**).

269 Behavioral state-specific OGs of COG category G included many phosphotransferase system
270 genes, in particular in Lactobacillus and Gilliamella species, but also glycoside hydrolases,
271 sugar binding and metabolism proteins. In most species, these OGs were more abundant in
272 foragers relative to nurses, except for Bifido_1.2 and Firm5_3, where the enrichment was in
273 the nurse samples, or spread equally between nurse and forager samples (**Supplementary**
274 **Dataset 1**).

275 The overrepresentation of the COG category H in Firm4_1 corresponded to 27 OGs involved
276 in the biosynthesis of cobalamin (vitamin B12) and adenosylcobalamin (coenzyme B12),
277 which had a higher coverage in forager than in nurse samples (as shown **Figure 3A**). In the
278 same species, we also found 18 genes for the utilization of ethanolamine among the OGs
279 enriched in foragers versus nurses. Processing of ethanolamine by the enzyme ammonia-
280 lyase is dependent on adenosylcobalamin (**Figure 3C**). To further explore the possible link
281 between the identified cobalamin biosynthesis and ethanolamine utilization OGs, we
282 analyzed the available isolate genomes of Firm4. Homologs of the cobalamin biosynthesis and

283 ethanol utilization pathways were only present in two (Hon2 and ESL0449) of the six Firm4_1
284 strains (**Figure 3D**) confirming that these genes are strain-specific. Moreover, the genes of
285 both pathways were encoded in the same genomic island, integrated into a conserved
286 chromosomal region providing further evidence that they are functionally linked and have
287 likely evolved together.

288 Another interesting set of behavioral state-specific OGs was identified in Gilli_1, Gilli_2, and
289 Snod_1, and encoded various components of Type VI secretion systems (T6SSs) involved in
290 interbacterial killing. This included genes annotated as effector proteins with PAAR or RHS
291 domains and immunity proteins. In Gilli_2, several genes encoding the structural components
292 of a T6SS entire secretion system were among the behavioral state-specific OGs. They
293 clustered in a genomic island which was present in only a single strain (A7) of Gilli_2 in our
294 reference database (**Supplementary Figure 8**). Strikingly, all these behavioral state-specific
295 T6SS-related OGs were more abundant in forager than in nurse samples.

296 In Snod_1, we also found other surface structure-related functions than T6SS among the
297 behavioral state-specific OGs, including genes encoding major subunits of Type IV pili, RTX
298 toxins, and proteins involved in capsular polysaccharide biosynthesis (**Supplementary**
299 **Dataset 1**).

300 ## Discussion

301 There is accumulating evidence that natural microbial communities harbor high extent of
302 strain-level diversity (Ansorge et al., 2019; Bendall et al., 2016; Tett et al., 2019; Zhu et al.,
303 2015). How such fine-scale diversity varies across space and time in nature is a key question
304 in microbial ecology and evolution (Goyal et al., 2022; Van Rossum et al., 2020; Zeevi et al.,
305 2019).

306 Here, we find that the strain composition of the honeybee gut microbiota undergoes a
307 consistent ecological turnover when bees transition from nurses to foragers. The polymorphic
308 site profiles of several community members clustered by behavioral state rather than by
309 colony or sampling location. Moreover, for most community members, we found consistent
310 differences in intraspecific gene content between nurses and foragers, which included
311 functions related to nutrient utilization, vitamin biosynthesis, and cell-cell interactions. In the
312 human gut microbiota, strain-level variation in individual community members have been
313 linked to diet (De Filippis et al., 2019) or health state (Carrow et al., 2020; Chen et al., 2021;
314 Greenblum et al., 2015; Zeevi et al., 2019). Our results show that a radical shift in host lifestyle
315 can change the strain composition of a large fraction of the gut microbiota, demonstrating
316 that strains represent ecological distinct units and emphasizing the importance to study
317 diversity in host-microbe associations at the strain level.

318 Previous shotgun metagenomics studies have focused on the analysis of individual
319 honeybees, and linked variation in species- and strain-level diversity in the gut microbiota to
320 host traits, such as age, colony origin, or genotype, but not to behavioral state (Ellegaard et
321 al., 2020; Ellegaard and Engel, 2019; Wu et al., 2021). In addition, we analyzed pooled samples
322 of 20 bees rather than individual bees, in order to include a larger number of colonies, from

323 different sites, and thereby increase the statistical power of the study. On the downside, the
324 pooling is probably responsible for the overall lower levels of polymorphic sites observed, as
325 it precludes the detection of polymorphisms present at low abundance or in a subset of the
326 sampled bee. However, rare strains have likely a minor impact on the whole community or
327 the host.

328 Another possible limitation of our study is that the sampled nurse and forager bees were not
329 age-matched. As foragers are usually older than nurses, age could have been a confounding
330 factor in our study. However, our previous findings have shown that the age of bees has little
331 impact on the strain level composition of the gut microbiota (Ellegaard et al. 2019). Therefore,
332 we are confident that most the observed differences are due to the nurse to forager transition
333 and the associated changes in dietary intake, physicochemical gut conditions, or host
334 physiology.

335 In contrast to nurses, foragers do not consume pollen. This is likely the main reason why
336 foragers were found to harbor lower bacterial biomass in the gut than nurses and exhibit a
337 consistent decrease of Bombilactobacillus, Lactobacillus, and Bifidobacterium relative to the
338 other phylotypes. This was already reported by two previous studies based on 16S rRNA gene
339 analysis (Jones et al., 2018; Kešnerová et al., 2020), one of which used age-matched forager
340 and nurse bees, corroborating the results of our shotgun metagenomics analyses. Pollen
341 constitutes an important nutrient source for bee gut-associated Bombilactobacillus,
342 Lactobacillus, and Bifidobacterium explaining their decrease in absolute abundance. All three
343 phylotypes harbor genes to degrade pollen-derived carbohydrates (Ellegaard et al., 2019;
344 Engel et al., 2012; Li et al., 2022; Zheng et al., 2019). and they readily grow in minimal media
345 supplement with pollen extract or entire pollen grains (Brochet et al., 2021; Kešnerová et al.,

346 2017). Moreover, experimentally colonized bees kept on a sugar water diet harbor
347 substantially fewer bacteria of these phylotypes in their gut than bees that also receive pollen
348 (Brochet et al., 2021; Kešnerová et al., 2020).

349 As previous studies were based on 16S rRNA gene sequencing, their analyses were limited to
350 the phylotype level. Hence, a key finding of our study is that not only phylotypes, but also
351 closely related species within phylotypes, and even different strains of the same species
352 responded differently to the nurse-to-forager transition. For example, one species of
353 Lactobacillus Firm5 barely changed in total abundance between foragers and nurses, while
354 others significantly decreased by >8 times. Likewise, some species showed consistent changes
355 in strain composition between nurses and foragers, while others did not. This suggests that
356 some species and strains are better than others in colonizing the gut environment of forager
357 bees, possibly because they are less dependent on pollen nutrients for growth, or are better
358 in withstanding changes in the physicochemical conditions in the gut as bees transition from
359 the nurse to the forager state. Furthermore, the reduced nutrient availability in the gut may
360 hinder strains to overgrow and dominate the community, explaining why some species
361 decreased in total abundance but increased in strain-level diversity in foragers.

362 If the observed shifts in strain composition were due to changes in the ecological niches in
363 the gut, we would expect to find functional differences between strains of the same species
364 dominating in nurses versus foragers. Previous studies have successfully identified such
365 condition-specific genes within host-associated bacterial species by comparing metagenomic
366 datasets. For example, strains of the human gut symbiont *Prevotella copri* harbored distinct
367 gene repertoires for drug metabolism and complex carbohydrate degradation in Western and
368 non-Western individuals (De Filippis et al., 2019). Likewise, strains of an intracellular symbiont

369 of *Bathymodiolus* deep-sea mussels showed substantial variation in gene content for nutrient
370 and energy acquisition indicating adaptation to the local conditions present in the
371 environment they were sampled from (Ansorge et al., 2019).

372 We detected strain-specific gene sets (i.e. OGs) associated with one of the two behavioral
373 states in almost all species of the gut microbiota. The fact that we found more behavioral
374 state specific OGs in foragers than in nurses can in part be explained by the higher extent of
375 strain-level diversity detected for some species in foragers, which logically also resulted in the
376 detection of more accessory gene content. Carbohydrate transport and metabolism functions
377 were prevalent among the behavioral state-specific gene content, which was somehow
378 expected as most species carry large accessory gene sets of this functional category in their
379 genomes (Ellegaard et al., 2019; Engel et al., 2012; Li et al., 2022). Moreover, previous studies
380 have shown that carbohydrate transport and metabolism are important for niche partitioning
381 of diet-derived carbohydrates in Lactobacillus and Bifidobacterium and that many of these
382 genes are differentially expressed when bees are fed pollen as compared to sugar water as
383 diet (Brochet et al., 2021; Zheng et al., 2019).

384 Somewhat more surprising was the finding that vitamin B12 biosynthesis and ethanolamine
385 utilization genes of Bombilactobacillus were specifically enriched in foragers versus nurses.
386 Some of these genes have already been found to be variably present among individual bees
387 in our previous study (Ellegaard and Engel 2019), but they could not be linked to any host
388 differences. Vitamin B12 is an important co-factor that cannot be synthesized by honeybees
389 (or other animals) (Kaval and Garsin, 2018), but needs to be taken up from the diet or
390 synthesized by bacterial symbionts (Elsayeh et al., 2022). It is possible that Bombilactobacillus
391 strains synthesizing this vitamin compensate the lack of vitamin B12 in the pollen-deprived

392 diet of foragers. However, a more plausible explanation for the increased abundance of these
393 genes in foragers may be that vitamin B12 is an essential co-factor for a major enzyme of the
394 ethanolamine utilization pathway (Kaval and Garsin, 2018). Ethanolamine is derived from the
395 membrane phospholipid phosphatidylethanolamine and constitutes an important bacterial
396 nitrogen and carbon source in the animal gut due to the high turnover of the host epithelium
397 and the microbiota (Tsoy et al., 2009). It is possible that ethanolamine accumulates in the
398 forager gut due to the presence of dead bacterial material or increased epithelial shedding,
399 providing a behavioral state-specific nutrient niche for the gut microbiota. The fact that the
400 genes for utilizing ethanolamine are strain-specific indicates that they are costly to maintain,
401 and that host state-specific nutritional niches may contribute to the maintenance of strain
402 level diversity in the gut microbiota.

403 We also found genes mediating bacteria-host interactions and interbacterial killing among
404 the behavioral state specific gene content. Of particular interest was the finding of several
405 T6SS genes associated with foragers in two of the three Gilliamella species and in
406 Snodgrassella. The primary niche of these bacteria is the ileum, where they form multispecies
407 biofilms on the host epithelium (Martinson et al., 2012). Diverse T6SS gene repertoires have
408 been identified in almost all sequenced strains of Gilliamella and Snodgrassella indicating that
409 they play key roles in interbacterial competition and niche establishment for these bacteria
410 (Steele et al., 2017; Steele and Moran, 2021). Our findings suggest that the forager gut
411 environment selects for strains with expanded capabilities in interbacterial warfare, possibly
412 as a consequence of the increased competition for nutrients and space. In Snodgrassella, we
413 also found genes encoding other surface structures (Type IV pili, LPS modifying enzymes, RTX
414 toxins) to be enriched in either nurses or foragers. Some of these genes are known
415 colonization factors of Snodgrassella (Powell et al., 2016), likely because they enable host

416 attachment and biofilm formation. However, why these functions are specifically enriched in
417 in one of the two behavioral states remains to be determined.

418 The possibility to carry out gnotobiotic bee experiments provides excellent opportunities for
419 future research to probe the functional relevance of the link between strain-specific gene
420 content and host state. Since the association of social bees with their microbiota is
421 evolutionary ancient, it will also be interesting to expand these analyses to closely related bee
422 species which show similar division of labor and whose microbiota has diversified to similar
423 extents as in *A. mellifera*.

424

425 ## Material and methods

426 ### Sampling

427 Adult female worker bees of *A. mellifera* were collected over the Summer 2019 from 15
428 different colonies in 5 different locations across Western Switzerland (**Supplementary Table**
429 **1**). For each sampled hive, we collected 20 larvae-feeding nurses and 20 pollen foragers, in
430 order to encompass most of the strain-level diversity within each colony. Samplings were
431 performed between 7 AM and 9 AM, when worker bees were beginning to return from their
432 foraging trips. Larvae-feeding nurses were identified as bees putting their heads in larvae
433 containing cells. Pollen foragers were identified as bees returning to their hive with pollen
434 pellets on their corbiculae. The bees were anesthetized using CO_2 and put on ice. They were
435 then dissected in sterile 1X PBS to extract the entire hindgut (pylorus, ileum, rectum) with
436 tools sterilized in 70% ethanol. Gut content was visually inspected to be consistent with life-
437 stage, i.e. nurse guts had to be pollen filled and forager guts had to be nectar-filled. Bees for
438 which the gut content could not be verified were excluded. Finally, the guts were pooled by
439 hive of origin and host phenotype (nurse and forager, respectively) (**Supplementary Table 1**).

440 ### DNA extraction

441 DNA extraction from the pooled gut samples was performed using the Qiagen Allprep
442 Powerfecal DNA/RNA extraction kit (Qiagen) following the kit handbook except for the bead-
443 beating steps. After dissection, the pooled guts were transferred into the kit's bead-beating
444 tubes. Lysis buffer and DDT were added following protocol instructions, and samples
445 underwent two bead-beating steps of 4 seconds at 6 m/s on a FastPrep-24 5G bead-beater
446 (MP Biomedicals). Gut samples with total weight over 220 mg were diluted to reach a
447 concentration of 220 mg in 650 µl lysis buffer + 25 µl DDT, then excess volume was removed

448 (i.e. for a gut weight of 440 mg, 337.5 µl were removed from the sample, and an identical
449 volume of lysis buffer + 3.8% DDT 1M was added to it). Subsequently, all samples underwent
450 a third bead-beating step for 45 seconds at 6 m/s. All the remaining steps were performed
451 following the manufacturer's protocol. DNA was diluted in 2x30 µl Buffer EB, its concentration
452 and quality were assessed with Qubit and Nanodrop, and samples stored at -20°C until further
453 processing.

454 **Quantitative real-time PCR**

455 We assessed the bacterial loads of the DNA samples using quantitative real-time PCR (qPCR)
456 following the methods of Ellegaard et al. (2020). We targeted both the V3-V4 region of the
457 16S rRNA bacterial gene and the actin gene of *A. mellifera* using the primers published by
458 Kešnerová et al (2017). We performed each qPCR reaction using a reaction of 1 µl extracted
459 DNA (10 ng/ µl), 3.2 µl nuclease-free water, 5 µl SYBR Select Master Mix (Life Technologies),
460 0.2 µl forward primers (5 µM) and 0.2 µl reverse primers (5 µM) in 96-well plates (MicroAmp
461 Fast 96-well Reaction Plate, Applied Biosystems for Life Technologies). The reactions were
462 performed in triplicates on a QuantStudio 5 Real-Time PCR System (Thermo Fischer Scientific)
463 using the parameters from Kešnerová et al (2017).

464 The data was extracted using the QuantStudio Design & Analysis Software v1.4 (Applied
465 Biosystems) and exported in csv format.

466 We computed the absolute number of each gene copy using the formula:

467 $$gene\ copy\ number = 10^{\frac{C_t - Intercept}{Slope}}$$

468 Intercept and slope parameters previously determined by our group (Kešnerová et al., 2017).

469 In order to compare the microbial gut loads of the different samples, we normalized the
470 number of 16S rRNA gene copies by the number of honeybee actin gene copies.
471 Normalization by the number of actin copies ensured we could compare the different samples
472 despite the different gut weights. The figures were generated using R (version 4.0) and the
473 ggplot2 library.

474 **Shotgun metagenomics sequencing**

475 The libraries were prepared with the Illumina Nextera Flex library kit (Illumina) with unique
476 dual indices (UDI) and sequenced on a HiSeq 4000 instrument (PE150) at the Genomic
477 Technologies Facility (GTF) of the University of Lausanne. All samples were multiplexed
478 together and sequenced on three lanes. The quality of the reads was assessed with FastQC
479 (v0.11.4, Babraham Institute) and subsequently trimmed and filtered to remove low quality
480 sequences and short reads using the tool Trimmomatic v0.35 (Bolger et al., 2014) with
481 settings PE ILLUMINACLIP:NexteraPE-PE.fa:2:30:10 TRAILING:20 SLIDINGWINDOW:4:20
482 MINLEN:50.

483 **Reference genome database**

484 The reference database of bacterial genomes was constructed in the same way as in Ellegaard
485 and Engel (2019) and Ellegaard et al. (2020). A total of 198 high-quality genomes of strains
486 isolated from the gut of honeybees and other social bee species such as bumble bees and
487 other *Apis* species were included in the database. Strains having more than 98.5% pairwise
488 ANI with other already included strains were not added to the database to limit redundancy.
489 The pairwise genomic ANI were computed using fastANI (Jain et al., 2018). In general, multiple
490 genomes were included for each bacterial phylotype. The database was annotated using the
491 tools eggnog-mapper (v1.0.3-40-g41a8498, Huerta-Cepas et al., 2019, 2017) and pfam (v1.6,

492 Bateman et al., 2002), from which BED files with gene positions were generated. Genes
493 smaller than 300 bp were filtered out. We predicted gene families of homologous genes using
494 OrthoFinder (v2.3.5, (Emms and Kelly, 2015)), and defined core genes as single-copy genes
495 present in all genomes of strains belonging to the same bacterial phylotype.

496 **Community composition analysis**

497 The community composition analysis was based on the work of Ellegaard and Engel (2019).
498 We mapped the reads using bwa mem (v0.7.15-r1142-dirty) (Li and Durbin, 2009) against the
499 reference genome database, and used samtools (v1.9) (Li et al., 2009) to remove chimeric
500 alignments with samtools view option –F 0x800. Alignments with less than 50 matches were
501 also removed. The read coverage information was extracted for each gene using the
502 command samtools bedcov and summed up across all orthologs of each core gene family for
503 genomes coming from the same species. Since genes located close to the origin of replication
504 had higher coverage than genes located close to the terminus (as an indication of replicating
505 bacteria (Bremer and Churchward, 1977)), the coverage at the terminus was inferred. To this
506 end, the coverage of each gene family was plotted relative to its genomic position using one
507 selected reference genome per species and a segmented regression line was fitted. The
508 relative abundance of each species in each sample was determined by dividing the terminus
509 coverage by the summed terminus coverage of all species, respectively. For species for which
510 the segmented lines did not show a proper fit (terminus location inferred far from estimated
511 breakpoint, or coverage at *Ori* lower than *Ter*), the median coverage of all core gene families
512 was used as an estimate for the relative abundance of that species.

513 **Strain-level analysis**

514 To quantify strain-level diversity within species we followed the approach established in
515 Ellegaard and Engel (2019). In short, we mapped the metagenomic reads using bwa mem to
516 a reduced version of the reference genome database including only one representative
517 member per species and used samtools to remove chimeric alignments. Reads mapped with
518 less than 50 matches registered in their CIGAR were also removed, as well as those whose
519 number of mismatches was higher than 5. Only species with more than 20x terminus coverage
520 were included. The bam files were subset for the core gene regions of each species, and genes
521 with less than 20x coverage were excluded. Positions within core genes having less than 10X
522 coverage in any of the samples were excluded from the analysis. Thus, for each species, the
523 same core gene positions were evaluated across all samples for determining polymorphic site
524 positions. To do so, we used the tool FreeBayes (v1.3.1-19-g54bf409) (Garrison and Marth,
525 2012) with options -—haplotype-length 0, --min-alternate-fraction 0.1, –-min-coverage 10, -C
526 5, -I, -X, -u and –-pooled-continuous. We used the tools vcfbreakmulti and vcfallelicprimitives
527 from the vcflib library (Garrison et al., 2021) to split complex observations into multiple lines.

528 Polymorphic sites for which the alternate fraction was lower than 0.1 for all samples were
529 filtered out (in order not to overestimate diversity in samples with deep coverage). We
530 computed the fraction of polymorphic sites for each bacterial species using the number of
531 identified sites divided by the length of the core genome. We also computed the cumulative
532 number of unique polymorphic sites across samples for each species using 10 random orders
533 of samples each time.

534 To assess the similarity between samples in terms of strain composition, we used the Jaccard
535 index based on the proportion of shared polymorphic sites; i.e. a polymorphic site where both

536 samples had a proportion of the alternative allele of at least 0.1 was considered shared
537 between the two samples. Non-polymorphic positions were filtered out. The distance
538 between all sample pairs was computed as $Distance(S1, S2) = 1 - Shared\ fraction(S1 -$
539 $S2)$. The matrix of pairwise distances between the samples was plotted using principal
540 coordinate analysis to identify samples that share similar polymorphic sites. We computed
541 the principal coordinates using the ape R library (v5.6-2).

542 **Functional gene content composition**

543 For the functional profiling of the gene content of the microbiota, we first extracted all the
544 reads which mapped to the bacterial database, plus those which did not map to the host
545 genome (*A. mellifera* genome (PRJNA471592, version Amel_Hav3.1), using the picard tools
546 (Broad Institute, 2019) (v2.7.1-SNAPSHOT) SamToFastq with parameters
547 VALIDATION_STRINGENCY=SILENT. We merged those reads for each sample and assembled
548 them using SPAdes (Nurk et al., 2017) with parameters --meta and -t 20. The obtained contigs
549 were filtered to keep only the ones with a length of at least 500 bp and a k-mer coverage of
550 at least 1.

551 We predicted the open-reading frame (ORF) positions using Prodigal (version 2.6.3, Hyatt et
552 al., 2010) with parameters -m -p meta. The ORFs with the flag partial=00 or length smaller
553 than 300 bp were filtered out. We annotated the remaining ORFs using eggNOG (v1.0.3-40-
554 g41a8498, Huerta-Cepas et al., 2019, 2017) with the script emapper.py and the options --
555 annotate_hits_table --no_file_comments --cpu 20.

556 To assign the contigs to the major bacterial species of the bee gut microbiota, we blasted the
557 predicted ORFs against a nucleotide database of all ORFs from our bacterial database using
558 the blastn tool (version 2.2.31+, Altschul et al., 1990) (evalue < 1e-05, pident > 50 and qcovs

559 > 50). Contigs were assigned to a given species when at least 80% of the ORFs were mapped
560 against genes of that species, as well as when at least 10% of the ORFs were mapped against
561 genes of that species and no mapping was found for any other species. We filtered out all
562 non-unanimous contigs (15198 contigs out of a total of 257015 (5.9%), representing 13.6% of
563 the bp).

564 For each species, we used the amino acid sequences from the ORFs and the genes from the
565 database strains to compute orthologous groups (OG) using OrthoFinder (Emms and Kelly,
566 2015). We created a database of all ORF's nucleotide sequences and mapped the reads
567 against it using bwa mem (version 0.7.15-r1142-dirty, Li and Durbin, 2009) . We removed
568 chimeric alignments using samtools view option –F 0x800 (version 1.9, Li et al., 2009).
569 Alignments with less than 50 matches were also removed. For each sample, we extracted the
570 coverage using the samtools coverage command (v1.15), and summed the coverage for the
571 ORFs belonging to the same OG. We normalized the coverage information for every sample
572 and every species using the OGs containing core genes: for each species of each sample, the
573 summed coverage of all core OGs was set to the same number (10000). To determine which
574 OGs were underrepresented or overrepresented in either the nurses or the foragers, we
575 computed the paired Wilcoxon sign rank test for each OG based on the normalized coverage
576 values, and then used qvalues to correct for multiple comparisons. We also computed the
577 fold-change by dividing the mean forager normalized coverage value by the nurse
578 counterpart. To be considered as significantly different between the treatments, an ORF had
579 to have both an absolute log2 fold change of at least 1, and a qvalue < 0.05. We finally
580 annotated the ORFs belonging to the significant OGs using the tool eggnog-mapper (1.0.3-40-
581 g41a8498, Huerta-Cepas et al., 2019, 2017). The whole analysis workflow is summarized in
582 **Supplementary Figure 7**.

Statistics and Data visualization

We performed the statistical analyses in R (v4.2.1) using the libraries ape (v5.6-2), data.table (v1.14.2), qvalue (v2.28.0), PERMANOVA (v0.2.0) and vegan (v2.6-2).

We generated all plots using R and the following R libraries: ggplot2 (v3.3.6), reshape (v0.8.9), gridExtra (v2.3), ggrepel (v0.9.1), RColorBrewer (v1.1-3) and scales (v1.2.1). Figures have been generated using Inkscape (v1.1).

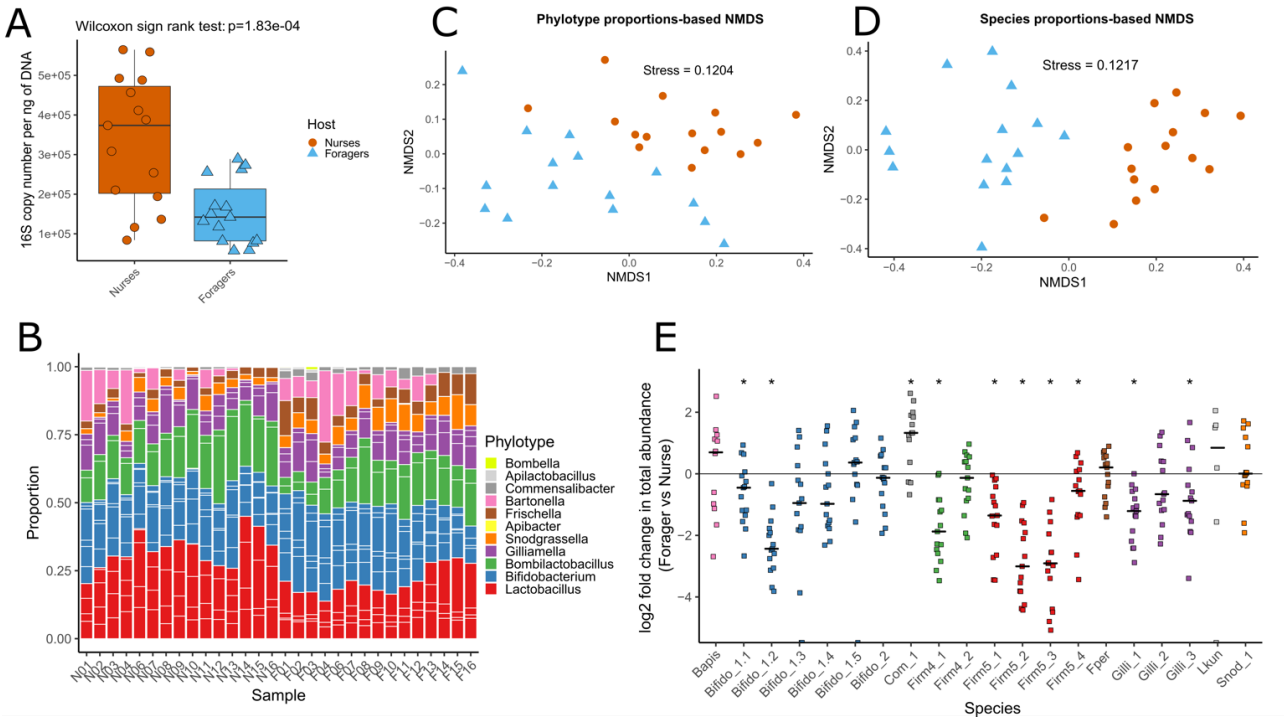
Data availability

The sequencing data is available under the BioProject ID PRJNA904667. The honeybee gut bacterial strain genome database is available at https://drive.switch.ch/index.php/s/vHmM6aIIFyDGQwm. The scripts required to run the analyses and produce the figures are available on GitHub (https://github.com/engelbeelab/Publication_Baud_metaG_NvsF_2022, commit f9be97f775b6bd41d1fcb003975e313e9617690e).

Acknowledgements

We would like to thank Mam Malick Sy Ndiaye, Aiswarya Prasad, Garance Sarton-Lohéac, Shinichi Sunagawa for providing helpful comments on the manuscript. We would like to thank Julien Marquis and the team from the Lausanne Genomics Technology Facility for carrying out the Illumina sequencing. We would like to thank Benjamin Dainat, Jan Dudda, Olivier Emery and Chloé Peillon for letting us sample bees from their apiaries. This work was funded by the University of Lausanne, an ERC Starting Grant (MicroBeeOme), the NCCR Microbiomes, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (grant no. 180575) and a Swiss National Science Foundation project grant (grant no. 179487) to P.E.

607 **Tables and figures**



608

609 **Figure 1.** Community composition analysis of the gut microbiota of the nurse and forager
610 worker bees. (A) Boxplot of the bacterial loads of each gut represented by the ratio between
611 the number of bacterial 16S rRNA gene copies and honeybee actin gene copies. (B) Stacked
612 bar plot of the bacterial phylotype composition for each sample. White separators within a
613 phylotype delineate the different species. (C) NMDS plot of the samples' bacterial phylotype
614 composition. (D) NMDS plot of the samples' bacterial species composition. (E) Log2 fold
615 change in total normalized absolute abundance of bacterial species in forager samples
616 relative to nurse samples from the same hive. Asterisks mark species with significant
617 differences between nurses and foragers (Wilcoxon sign rank test of the normalized absolute
618 abundance of each species in the foragers against the nurses, qvalue < 0.05). Data points for
619 species of the same phylotype are shown in the same color. Horizontal bars represent
620 medians.

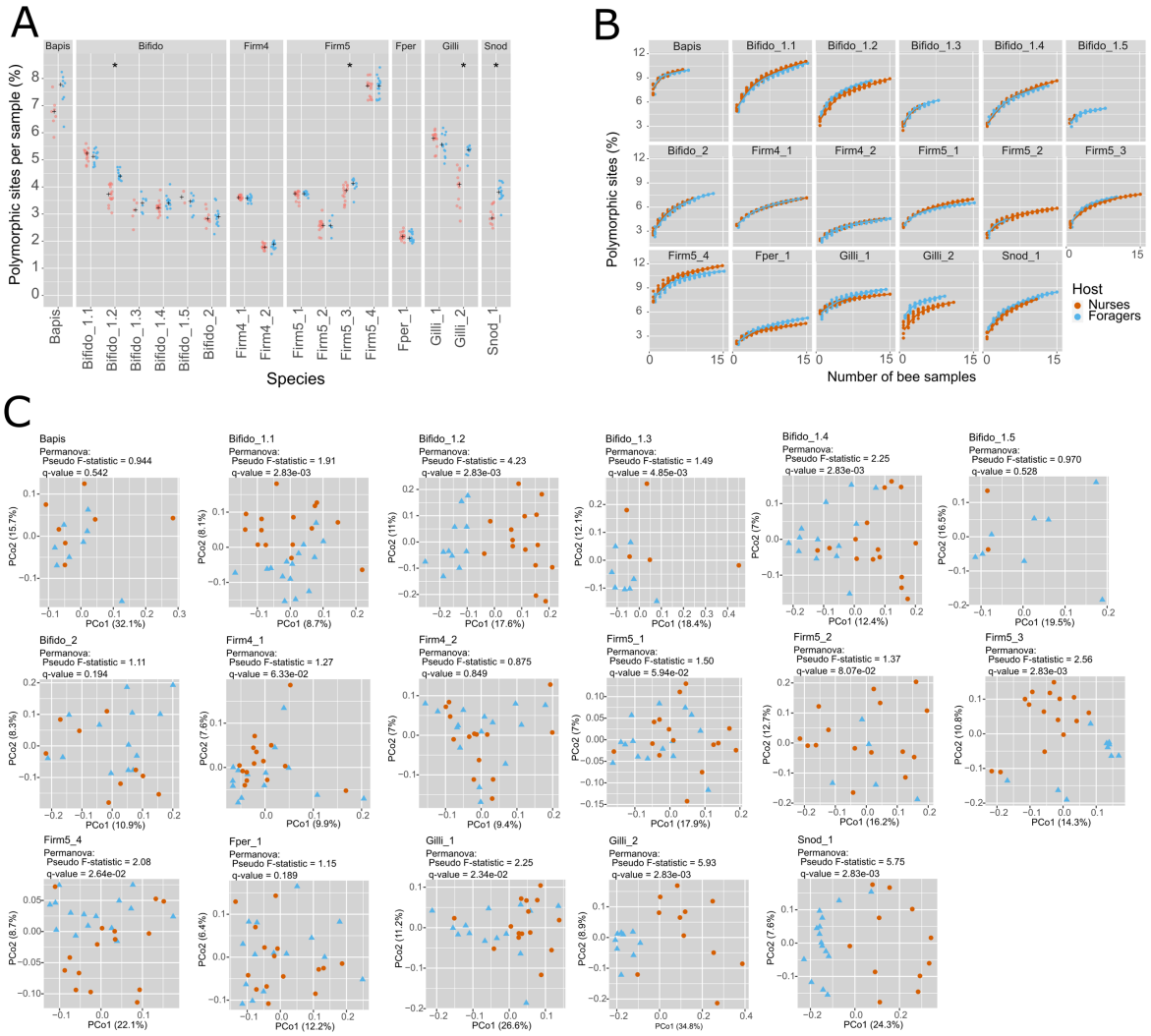


Figure 2. Differences in strain-level diversity between nurse and forager samples. (A) Proportion of polymorphic sites in the core genes for the different species, expressed as a percentage of the total core genome length. Asterisks symbolize a qvalue < 0.05 for the Wilcoxon sign rank test between the proportion of polymorphic sites between nurse and forager samples for the species. (B) Curves of the cumulative proportions of polymorphic sites for the different species. (C) Principal Coordinate Analysis plots based on the Jaccard distance between each pair of samples for each of the species.

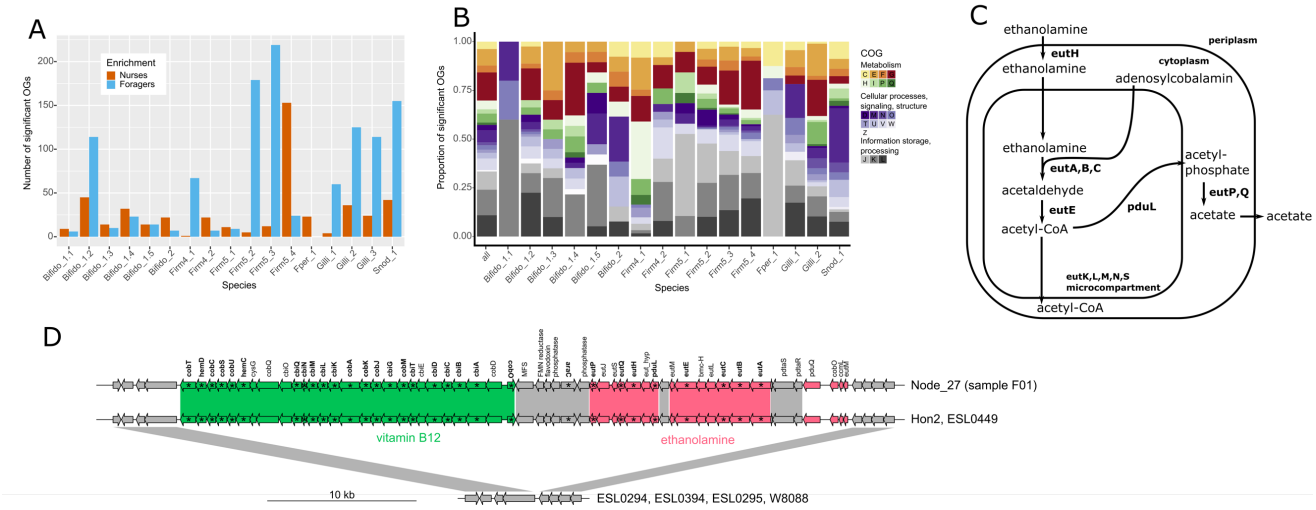
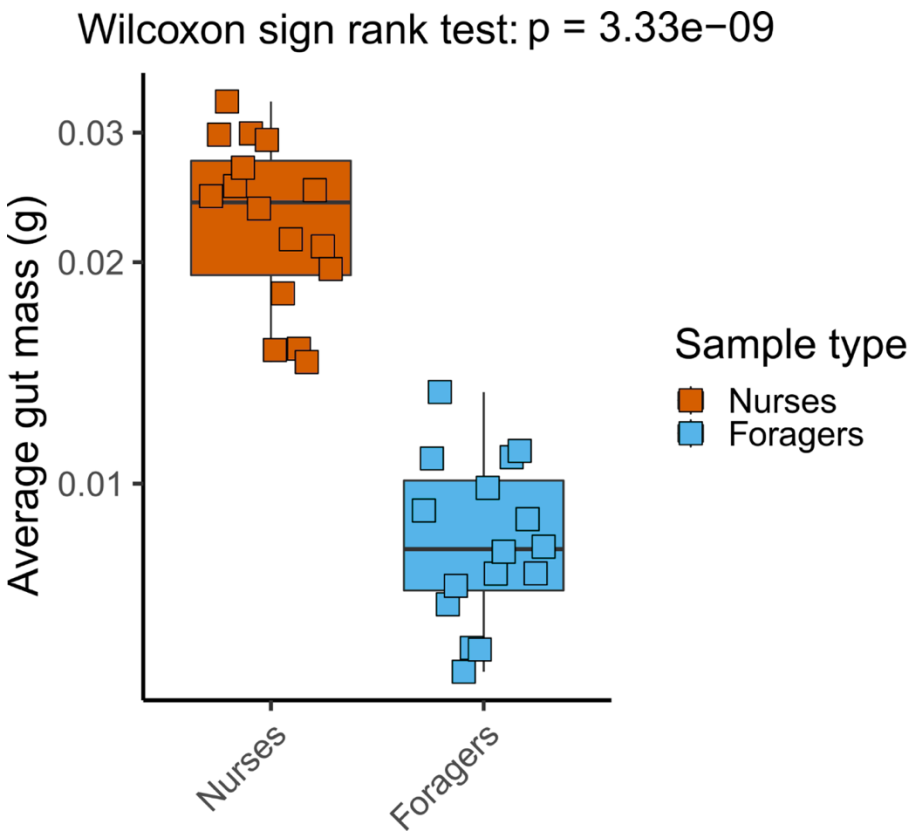


Figure 3. Differences in functional gene content within species of the gut microbiota between nurse and forager samples. (A) Number of orthologous groups (OGs) significantly differing in coverage for each of the species. (B) Repartition of the COG categories of the significant OGs, nurse-enriched and forager-enriched included. OGs that were not categorized or categorized as COG S were not included in the plot. (C) Figure adapted from (Garsin, 2010). Possible ethanolamine utilization pathway enabled by the operon shown in (D). (D) Genomic island containing the ethanolamine utilization and vitamin B12 biosynthesis genes, and its presence in the isolate genomes included in our bacterial reference database. Genes enriched in the forager samples are in bold and flagged with an asterisk.

642 **Supplementary tables and figures**

643

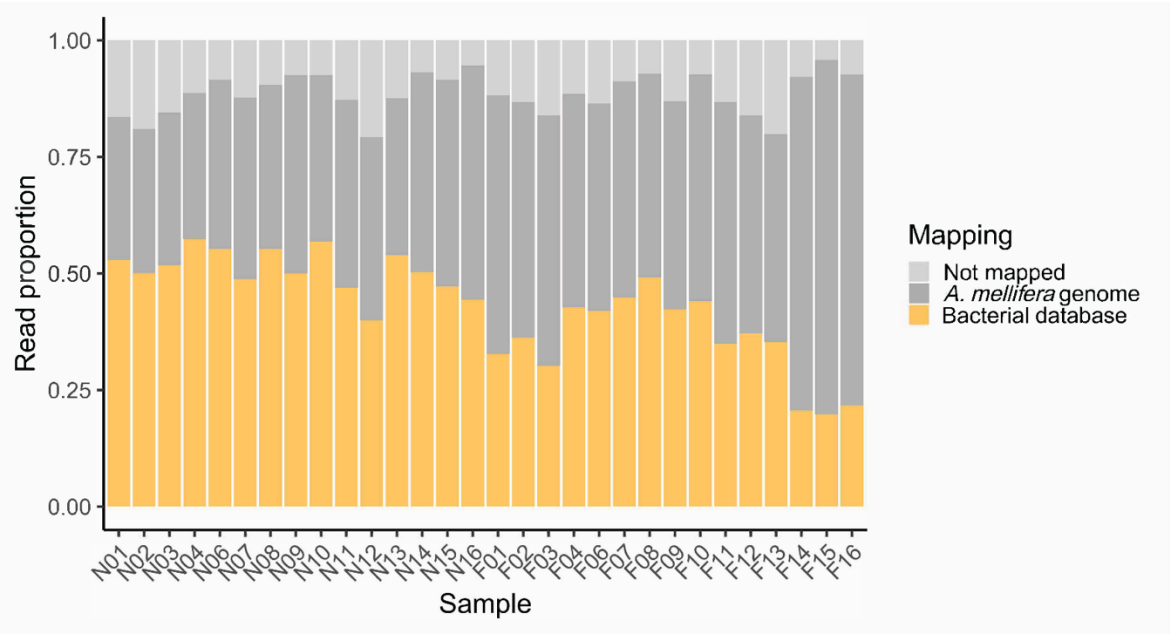


644

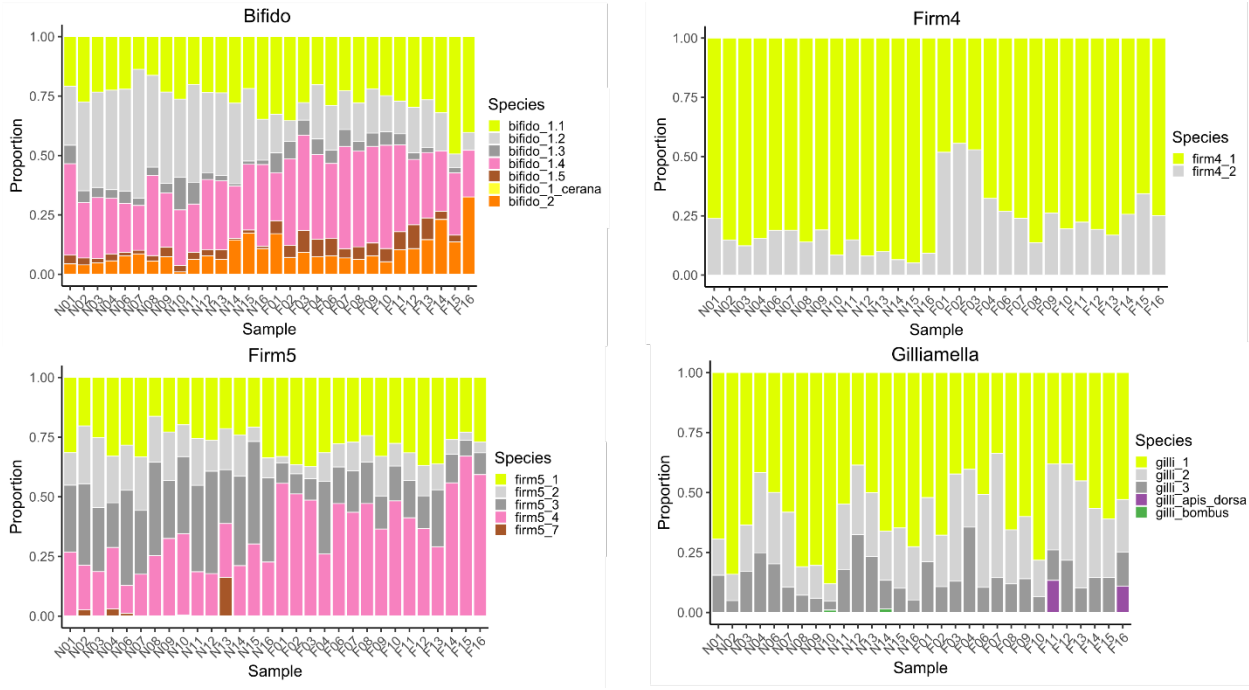
645 **Supplementary Figure 1.** Per sample average gut mass, computed as each sample's gut pool

646 mass, divided by the number of guts in the pool.

647

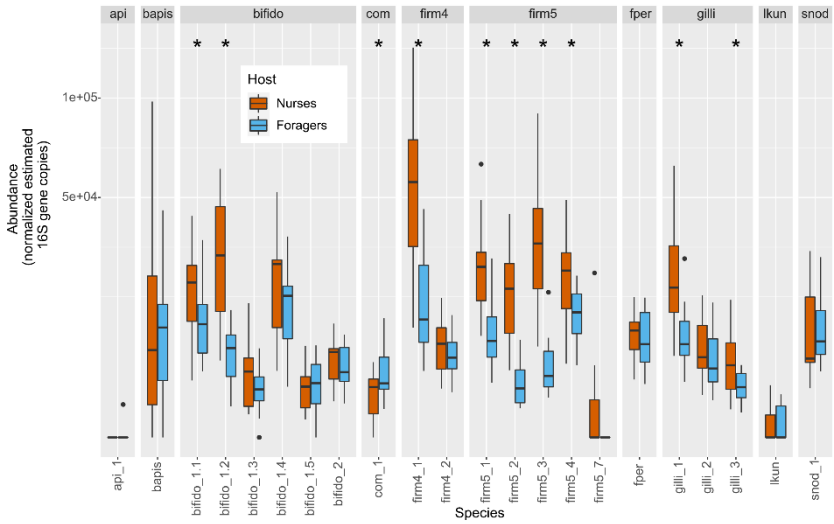


Supplementary Figure 2. Stacked bar plot of the proportions of reads mapping to the bacterial database and to the *A. mellifera* genome, as well as non-mapped reads, for each of the samples.



Supplementary Figure 3. Stacked bar plots of the proportions of the different species detected for the multi-species phylotypes.

658



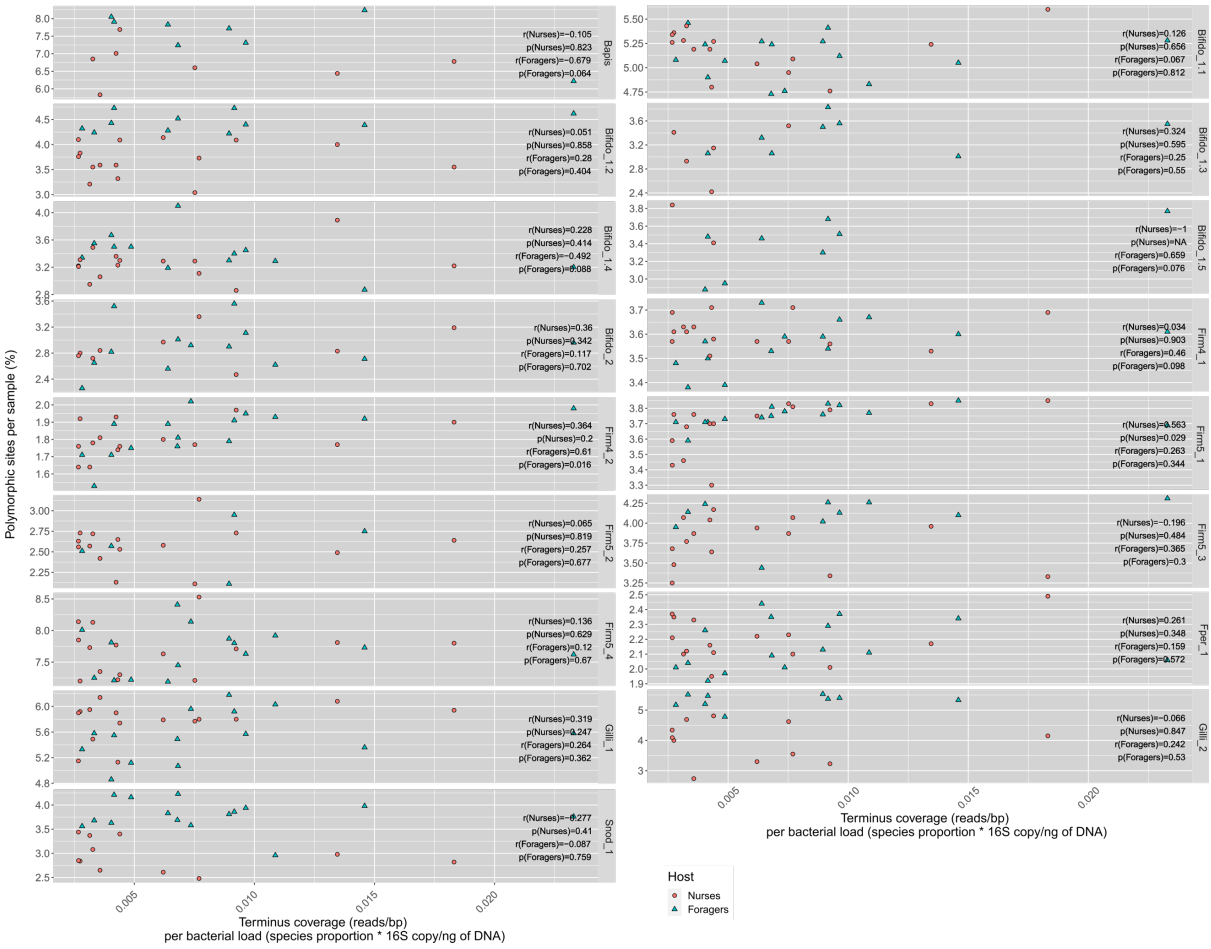
659

660 **Supplementary Figure 4.** Absolute abundance of the different species detected in the

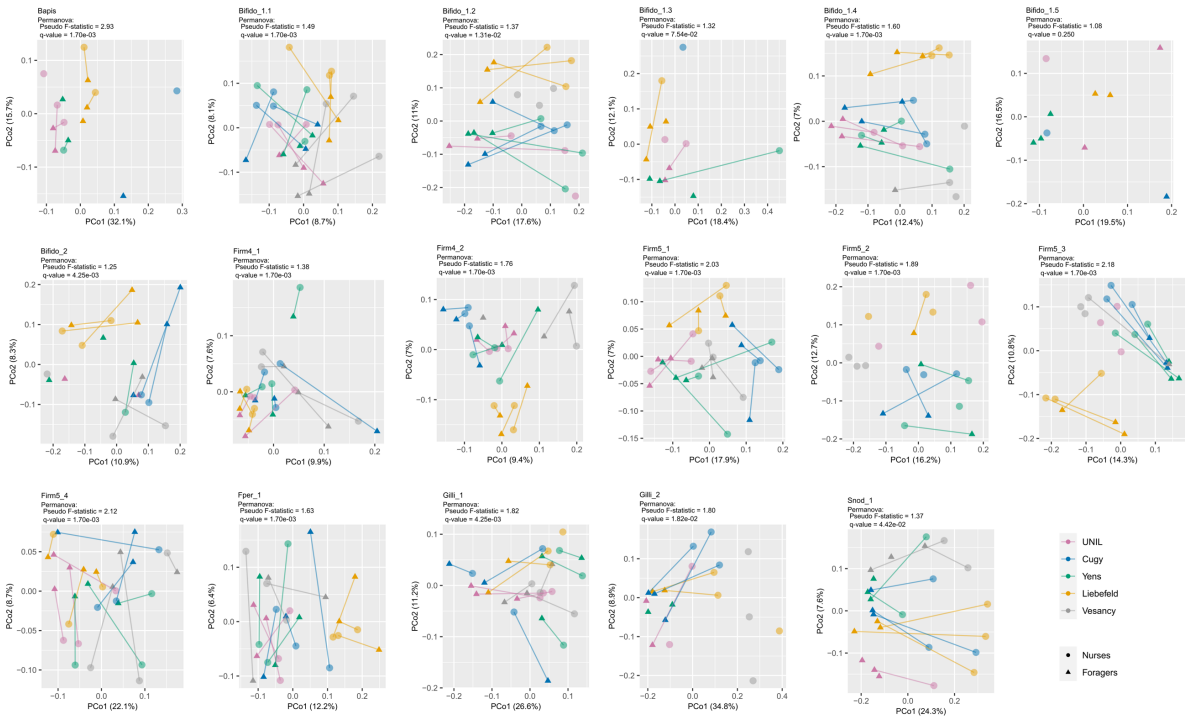
661 samples approximated by the normalized estimated number of 16S rRNA gene copies.

662 Asterisks indicate q-values < 0.05.

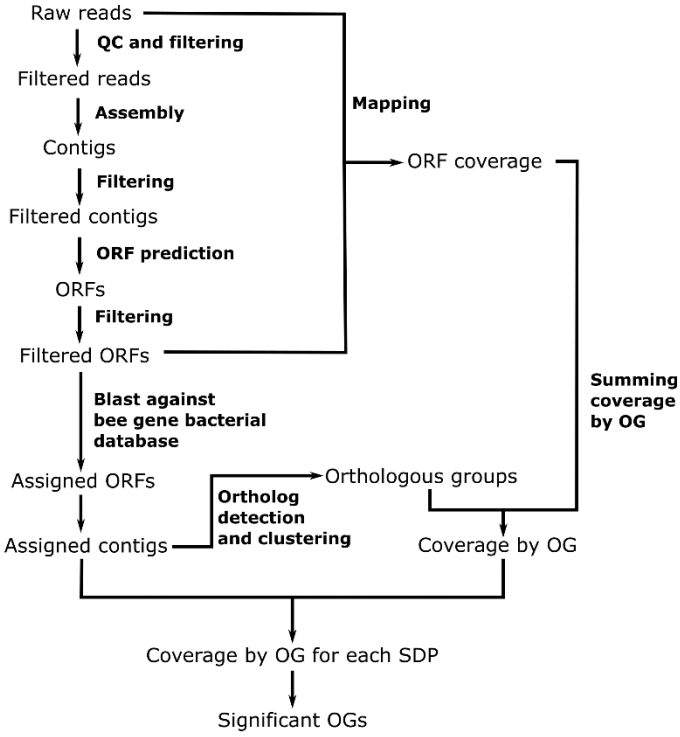
663



Supplementary Figure 5. Pearson correlation coefficient and non-corrected p-value for the correlation of the percentage of polymorphic sites per species with the coverage of the species normalized by the bacterial load for each sample.



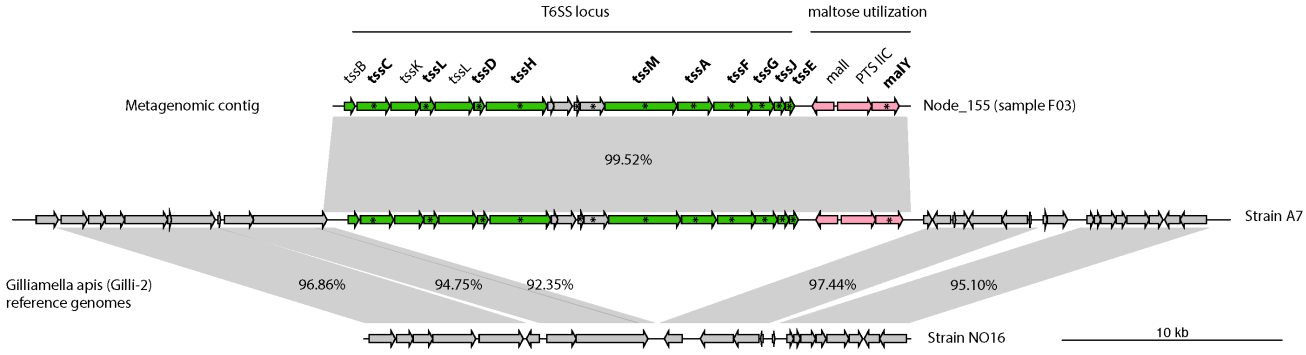
Supplementary Figure 6. Principal Coordinate Analyses of the Jaccard distance between all samples for the different species, based on presence or absence of the polymorphic sites. The color scheme shows information on the sampling sites of the different samples. The connecting lines indicates the two samples from the same colony.



676

677 **Supplementary Figure 7.** Workflow used for the functional gene content analysis.

678



Supplementary Figure 8. Genomic island of Gilli_2 encoding T6SS genes enriched in forager versus nurse samples. The genomic island is only present in strain A7 of Gilli_2 in our reference database and is integrated in a conserved genomic region. Genes in green indicate T6SS genes, genes in pink depict strain-specific genes beloning to the same genomic island, and genes in grey are other other genes present in the flanking regions of genomes of Gilli_2. Genes enriched in forager samples are labeled in bold and flagged with an asterisk.

687 **Supplementary Table 1.** Sample collection information.

Sample name	Collection date	Treatment	Location	Number of guts
N01	25.06.2019	Larvae-feeding nurses	Lausanne	19
N02	25.06.2019	Larvae-feeding nurses	Lausanne	18
N03	25.06.2019	Larvae-feeding nurses	Lausanne	20
N04	26.07.2019	Larvae-feeding nurses	Liebefeld	20
N06	26.07.2019	Larvae-feeding nurses	Liebefeld	20
N07	26.07.2019	Larvae-feeding nurses	Liebefeld	20
N08	30.07.2019	Larvae-feeding nurses	Yens	20
N09	30.07.2019	Larvae-feeding nurses	Yens	20
N10	30.07.2019	Larvae-feeding nurses	Yens	20
N11	06.08.2019	Larvae-feeding nurses	Cugy	20
N12	06.08.2019	Larvae-feeding nurses	Cugy	20
N13	06.08.2019	Larvae-feeding nurses	Cugy	20
N14	27.08.2019	Larvae-feeding nurses	Vesancy	20
N15	27.08.2019	Larvae-feeding nurses	Vesancy	20
N16	27.08.2019	Larvae-feeding nurses	Vesancy	20
F01	25.06.2019	Pollen foragers	Lausanne	20
F02	25.06.2019	Pollen foragers	Lausanne	20
F03	25.06.2019	Pollen foragers	Lausanne	20
F04	26.07.2019	Pollen foragers	Liebefeld	20
F06	26.07.2019	Pollen foragers	Liebefeld	20
F07	26.07.2019	Pollen foragers	Liebefeld	20
F08	30.07.2019	Pollen foragers	Yens	20
F09	30.07.2019	Pollen foragers	Yens	20

F10	30.07.2019	Pollen foragers	Yens	20
F11	06.08.2019	Pollen foragers	Cugy	11
F12	06.08.2019	Pollen foragers	Cugy	16
F13	06.08.2019	Pollen foragers	Cugy	10
F14	27.08.2019	Pollen foragers	Vesancy	20
F15	27.08.2019	Pollen foragers	Vesancy	20
F16	27.08.2019	Pollen foragers	Vesancy	20

690 **Supplementary Table 2.** Table of all strain genomes included in the reference database, their
691 Species/SDP abbreviation used in this study, their corresponding full species name (if existing)
692 and the host where they were isolated.

Locus_tag	Strain_name	Phylotype	Species/SDP abbreviation	Species	Host
C4S75	wkB309	*Apibacter*	Api_1	*Apibacter sp.*	Apis cerana
Ga0307799	ESL0404	*Apibacter*	Api_1	*Apibacter sp.*	Apis cerana
Ga0312307	ESL0432	*Apibacter*	Api_1	*Apibacter sp.*	Apis cerana
C4S76	wkB180	*Apibacter*	Api_apis_dorsa	*Apibacter adventoris*	Apis dorsata
C4S77	wkB301	*Apibacter*	Api_apis_dorsa	*Apibacter adventoris*	Apis dorsata
Ga0061079	R-53146	*Apibacter*	Api_bombus	*Apibacter mensalis*	Bombus lapidarius
BBC0122	BBC0122	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
BBC0178	BBC0178	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
BBC0244	BBC0244	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S82	M0192	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S80	M0177	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S83	W8122	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S84	W8098	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S85	M0187	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S89	B10834G6	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S90	W8097	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3S98	B10834H15	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3U94	W8125	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3V12	W8167	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3V13	M0280	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
H3V16	W8099	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera

H3V17	M0283	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
PEB0122	PEB0122	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
PEB0149	PEB0149	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
PEB0150	PEB0150	*Bartonella*	Bapis	*Bartonella apis*	Apis mellifera
Ga0072395	Bin2	*Bifidobacterium*	Bifido_1.1	*bifidobacterium apousia*	Apis mellifera
H3T81	M0109	*Bifidobacterium*	Bifido_1.1	*bifidobacterium apousia*	Apis mellifera
H3T92	W8112	*Bifidobacterium*	Bifido_1.1	*bifidobacterium apousia*	Apis mellifera
H3U88	W8120	*Bifidobacterium*	Bifido_1.1	*bifidobacterium apousia*	Apis mellifera
Ga0072396	Bin7	*Bifidobacterium*	Bifido_1.2	*Bifidobacterium choladohabitans*	Apis mellifera
H3S97	B14384H11	*Bifidobacterium*	Bifido_1.2	*Bifidobacterium choladohabitans*	Apis mellifera
H3T86	W8113	*Bifidobacterium*	Bifido_1.2	*Bifidobacterium choladohabitans*	Apis mellifera
H3U87	W8101	*Bifidobacterium*	Bifido_1.2	*Bifidobacterium choladohabitans*	Apis mellifera
H3U89	W8104	*Bifidobacterium*	Bifido_1.2	*Bifidobacterium choladohabitans*	Apis mellifera
H3U98	W8116	*Bifidobacterium*	Bifido_1.2	*Bifidobacterium choladohabitans*	Apis mellifera
H3V04	M0353	*Bifidobacterium*	Bifido_1.2	*Bifidobacterium choladohabitans*	Apis mellifera
BAST	PRL2011	*Bifidobacterium*	Bifido_1.3	*Bifidobacterium asteroides*	Apis mellifera
Ga0133549	ESL0170	*Bifidobacterium*	Bifido_1.3	*Bifidobacterium asteroides*	Apis mellifera
H3S92	W8130	*Bifidobacterium*	Bifido_1.3	*Bifidobacterium asteroides*	Apis mellifera
H3S93	W8109	*Bifidobacterium*	Bifido_1.3	*Bifidobacterium asteroides*	Apis mellifera
H3T89	W8114	*Bifidobacterium*	Bifido_1.3	*Bifidobacterium asteroides*	Apis mellifera

Ga0072401	Hma3	*Bifidobacterium*	Bifido_1.4	*Bifidobacterium polysaccharolyticum*	Apis mellifera
Ga0133551	ESL0198	*Bifidobacterium*	Bifido_1.4	*Bifidobacterium polysaccharolyticum*	Apis mellifera
H3T90	W8117	*Bifidobacterium*	Bifido_1.4	*Bifidobacterium polysaccharolyticum*	Apis mellifera
H3T91	B14448H7	*Bifidobacterium*	Bifido_1.4	*Bifidobacterium polysaccharolyticum*	Apis mellifera
H3T93	M0404	*Bifidobacterium*	Bifido_1.4	*Bifidobacterium polysaccharolyticum*	Apis mellifera
H3V06	W6238G3	*Bifidobacterium*	Bifido_1.4	*Bifidobacterium polysaccharolyticum*	Apis mellifera
Ga0133552	ESL0199	*Bifidobacterium*	Bifido_1.4	*Bifidobacterium sp.*	Apis mellifera
Ga0133553	ESL0200	*Bifidobacterium*	Bifido_1.5	*Bifidobacterium sp.*	Apis mellifera
Ga0133550	ESL0197	*Bifidobacterium*	Bifido_2	*Bifidobacterium indicum/coryneforme*	Apis mellifera
Ga0072398	Bma6	*Bifidobacterium*	Bifido_2	*Bifidobacterium indicum/coryneforme*	Apis mellifera
BINDI	LMG 11587	*Bifidobacterium*	Bifido_2	*Bifidobacterium indicum*	Insect
BCOR	LMG 18911	*Bifidobacterium*	Bifido_2	*Bifidobacterium coryneforme*	Apis mellifera
H3V08	M0307	*Bifidobacterium*	Bifido_2	*Bifidobacterium sp.*	Apis mellifera
Ga0098206	DSM 22766	*Bifidobacterium*	Bifido_bombus	*Bifidobacterium actinocoloniiforme*	Bumblebee digestive tract
Ga0056940	DSM 19703	*Bifidobacterium*	Bifido_bombus	*Bifidobacterium bombi*	Bumblebee digestive tract
Ga0057557	DSM 22767	*Bifidobacterium*	Bifido_bombus	*Bifidobacterium bohemicum*	Bumblebee digestive tract

Ga0312305	ESL0447	*Bifidobacterium*	Bifido_1_cerana	*Bifidobacterium sp.*	Apis cerana
Ga0372762	ESL_0368	*Bombella*	Bom_1	*Bombella sp.*	Apis mellifera
Ga0216357	ESL_0385	*Bombella*	Bom_apis_melli	*Bombella sp.*	Apis mellifera
Ga0308518	R-52487	*Bombella*	Bom_bombus	*Bombella intestini*	Bombus lapidarius
D9V35	AMU001	*Commensalibacter*	Com_1	*Commensalibacter*	Apis mellifera
H3T57	W6292M3	*Commensalibacter*	Com_1	*Commensalibacter*	Apis mellifera
CIN	A911	*Commensalibacter*	Com_drosophila	*Commensalibacter intestini*	Drosophila melagenogaster
Ga0248239	MX01	*Commensalibacter*	Com_monarch	*Commensalibacter*	Monarch butterfly
Ga0070888	Hon2	*Bombilactobacillus*	Firm4_1	*Bombilactobacillus mellis*	Apis mellifera
Ga0227357	ESL0294	*Bombilactobacillus*	Firm4_1	*Bombilactobacillus mellis*	Apis mellifera
Ga0227358	ESL0295	*Bombilactobacillus*	Firm4_1	*Bombilactobacillus mellis*	Apis mellifera
Ga0227360	ESL0394	*Bombilactobacillus*	Firm4_1	*Bombilactobacillus mellis*	Apis mellifera
Ga0326456	ESL0449	*Bombilactobacillus*	Firm4_1	*Bombilactobacillus mellis*	Apis cerana
H3T38	W8088	*Bombilactobacillus*	Firm4_1	*Bombilactobacillus mellis*	Apis mellifera
Ga0072399	Bin4	*Bombilactobacillus*	Firm4_2	*Bombilactobacillus mellifer*	Apis mellifera

H3R26	W8092	*Bombilactobacillus*	Firm4_2	*Bombilactobacillus mellifer*	Apis mellifera
Ga0133563	ESL0185	*Lactobacillus*	Firm5_1	*Lactobacillus apis*	Apis mellifera
Ga0061073	LMG 26964	*Lactobacillus*	Firm5_1	*Lactobacillus apis*	Apis mellifera
Ga0072402	ESL0353	*Lactobacillus*	Firm5_1	*Lactobacillus apis*	Apis mellifera
Ga0226842	ESL0263	*Lactobacillus*	Firm5_1	*Lactobacillus apis*	Apis mellifera
H3R21	W8172	*Lactobacillus*	Firm5_1	*Lactobacillus apis*	Apis mellifera
H3R25	M0390	*Lactobacillus*	Firm5_1	*Lactobacillus apis*	Apis mellifera
H3T48	M0403	*Lactobacillus*	Firm5_1	*Lactobacillus apis*	Apis mellifera
Ga0133561	ESL0183	*Lactobacillus*	Firm5_2	*Lactobacillus helsingborgensis*	Apis mellifera
Ga0072400	ESL0354	*Lactobacillus*	Firm5_2	*Lactobacillus helsingborgensis*	Apis mellifera
LACWKB8	wkB8	*Lactobacillus*	Firm5_2	*Lactobacillus helsingborgensis*	Apis mellifera
H3T47	W8093	*Lactobacillus*	Firm5_2	*Lactobacillus helsingborgensis*	Apis mellifera
Ga0133562	ESL0184	*Lactobacillus*	Firm5_3	*Lactobacillus melliventris*	Apis mellifera
Ga0072404	ESL0350	*Lactobacillus*	Firm5_3	*Lactobacillus melliventris*	Apis mellifera
Ga0225908	ESL0259	*Lactobacillus*	Firm5_3	*Lactobacillus melliventris*	Apis mellifera
Ga0226840	ESL0260	*Lactobacillus*	Firm5_3	*Lactobacillus melliventris*	Apis mellifera
Ga0227359	ESL0393	*Lactobacillus*	Firm5_3	*Lactobacillus melliventris*	Apis mellifera
H3R22	W8171	*Lactobacillus*	Firm5_3	*Lactobacillus melliventris*	Apis mellifera
Ga0072403	ESL0352	*Lactobacillus*	Firm5_4	*Lactobacillus kimbladii*	Apis mellifera
Ga0070887	ESL0351	*Lactobacillus*	Firm5_4	*Lactobacillus kullabergensis*	Apis mellifera
LACWKB10	wkB10	*Lactobacillus*	Firm5_4	*Lactobacillus kullabergensis*	Apis mellifera

Ga0226839	ESL0261	*Lactobacillus*	Firm5_4	*Lactobacillus kullabergensis*	Apis mellifera
Ga0133564	ESL0186	*Lactobacillus*	Firm5_4	*Lactobacillus kullabergensis*	Apis mellifera
H3T43	M0396	*Lactobacillus*	Firm5_4	*Lactobacillus kullabergensis*	Apis mellifera
H3U49	W8174	*Lactobacillus*	Firm5_4	*Lactobacillus kullabergensis*	Apis mellifera
Ga0307804	ESL0409	*Lactobacillus*	Firm5_7	*Lactobacillus panisapium*	Apis cerana
Ga0311394	ESL0411	*Lactobacillus*	Firm5_7	*Lactobacillus panisapium*	Apis cerana
Ga0312303	ESL0413	*Lactobacillus*	Firm5_7	*Lactobacillus panisapium*	Apis cerana
Ga0312304	ESL0416	*Lactobacillus*	Firm5_7	*Lactobacillus panisapium*	Apis cerana
Ga0226852	ESL0228	*Lactobacillus*	Firm5_bombus_1	*Lactobacillus bombicola*	Bombus bohemicus
Ga0226843	ESL0233	*Lactobacillus*	Firm5_bombus_2	*Lactobacillus bombicola*	Bombus terrestris
Ga0226847	ESL0230	*Lactobacillus*	Firm5_bombus_2	*Lactobacillus bombicola*	Bombus pascuorum
Ga0061072	R-53102	*Lactobacillus*	Firm5_bombus_2	*Lactobacillus bombicola*	Bombus lapidarius
Ga0077910	ESL0034	*Frischella*	Fper_1	*Frischella perrara*	Apis mellifera
Ga0133554	ESL0167	*Frischella*	Fper_1	*Frischella perrara*	Apis mellifera
Ga0418777	Fper_2	*Frischella*	Fper_2	*Frischella sp.*	Apis cerana
Ga0133559	ESL0178	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
GAPWK	wkB1	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
B6D22	AM4	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
A9G14	P54G	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
A9G17	wkB7	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera

B6D03	N-G5	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
B6D19	N-22	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
B6D08	A-7-12	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
B5S40	A-1-24	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
H3S73	B14384G12	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
H3T96	W8136	*Gilliamella*	Gilli_1	*Gilliamella apicola*	Apis mellifera
Ga0133555	ESL0169	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
Ga0133556	ESL0172	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
B6C99	N-G2	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
B6D18	A7	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
A9G16	P83G	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
A9G19	M1-2G	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
A9G15	P62G	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
B6C96	NO15	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
B6D14	N-G4	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
B6D17	N-G1	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
B6D21	AM1	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
B6C88	NO16	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
H3U96	M0364	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
H3T80	W8145	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
H3T56	W8123	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
H3T83	M0320	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
H3U97	W8128	*Gilliamella*	Gilli_2	*Gilliamella apis*	Apis mellifera
Ga0133557	ESL0177	*Gilliamella*	Gilli_3	*Gilliamella apis*	Apis mellifera
Ga0133558	ESL0182	*Gilliamella*	Gilli_3	*Gilliamella apis*	Apis mellifera
Ga0307802	ESL0441	*Gilliamella*	Gilli_4	*Gilliamella sp.*	Apis cerana
A9G10	wkB308	*Gilliamella*	Gilli_4	*Gilliamella sp.*	Apis cerana

Ga0307800	ESL0405	*Gilliamella*	Gilli_5	*Gilliamella sp.*	Apis cerana
Ga0307803	ESL0443	*Gilliamella*	Gilli_6	*Gilliamella sp.*	Apis cerana
A9G07	wkB72	*Gilliamella*	Gilli_6	*Gilliamella sp.*	Apis cerana
A9G08	wkB195	*Gilliamella*	Gilli_6	*Gilliamella sp.*	Apis cerana
A9G09	wkB292	*Gilliamella*	Gilli_apis_andre	*Gilliamella sp.*	Apis cerana
A9G03	wkB171	*Gilliamella*	Gilli_apis_andre	*Gilliamella sp.*	Apis andreniformis
A9G11	wkB108	*Gilliamella*	Gilli_apis_dorsa	*Gilliamella sp.*	Apis dorsata
A9G12	wkB112	*Gilliamella*	Gilli_apis_dorsa	*Gilliamella sp.*	Apis dorsata
A9G13	wkB178	*Gilliamella*	Gilli_apis_dorsa	*Gilliamella sp.*	Apis dorsata
A9G23	App4-10	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus appositus
A9G24	App6-5	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus appositus
A9G29	Fer2-1	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus fervidus
Ga0055041	wkB11	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus bimaculatus
A9G30	Fer4-1	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus fervidus
A9G31	Gris1-4	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus griseocollis
A9G32	Gris3-2	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus griseocollis
A9G35	Choc5-1	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus impatiens
A9G37	GilliExp13	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus impatiens
A9G38	Imp1-1	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus impatiens
A9G39	Imp1-6	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus impatiens

A9G41	Nev5-1	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus nevadensis
A9G42	Nev6-6	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus nevadensis
A9G43	Occ3-1	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus occidentalis
A9G44	Occ4-3	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus occidentalis
A9G47	WF3-4	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus pensylvanicus
Ga0227298	ESL0232	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus terrestris
Ga0227301	ESL0250	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus pascuorum
Ga0227302	ESL0254	*Gilliamella*	Gilli_bombus	*Gilliamella sp.*	Bombus pascuorum
Ga0061080	R-53144	*Gilliamella*	Gilli_bombus	*Gilliamella intestini*	Soil
Ga0098617	Fhon2	*Apilactobacillus*	Lkun	*Apilactobacillus kunkeei*	Apis mellifera
Ga0308523	Fhon13	*Apilactobacillus*	Lkun	*Apilactobacillus kunkeei*	Apis mellifera
H3T37	M0345	*Apilactobacillus*	Lkun	*Apilactobacillus kunkeei*	Apis mellifera
Ga0326500	ESL0196	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
SALWKB2	wkB2	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
CPT77	E1	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGH94	Aw-20	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGH95	Aw-18	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGH96	A12	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGH97	A-10-12	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGI00	A-1-12	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGI01	A2	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera

BGI08	N-W7	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGI09	N9	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGI12	N-S3	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BGI13	N-S4	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BHC42	wkB332	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BHC47	PEB0171	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BHC50	wkB339	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BHC52	MS1-3	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
Ga0227305	ESL0323	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
Ga0227306	ESL0324	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
Ga0227304	ESL0304	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
H3U73	W8124	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
H3T75	W8135	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
H3T79	M0118	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
H3U82	W8132	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
H3V10	W6238H14	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
H3V01	M0351	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
H3V11	W8158	*Snodgrassella*	Snod_1	*Snodgrassella alvi*	Apis mellifera
BHC53	wkB237	*Snodgrassella*	Snod_2	*Snodgrassella sp.*	Apis andreniformis
BHC43	wkB298	*Snodgrassella*	Snod_2	*Snodgrassella sp.*	Apis cerana
BHC54	wkB273	*Snodgrassella*	Snod_2	*Snodgrassella sp.*	Apis florea
BHC44	Nev4-2	*Snodgrassella*	Snod_bombus	*Snodgrassella sp.*	Bombus nevadensis
BHC46	Ruf1-X	*Snodgrassella*	Snod_bombus	*Snodgrassella alvi*	Bombus rufocinctus
BHC48	Occ4-2	*Snodgrassella*	Snod_bombus	*Snodgrassella communis*	Bombus occidentalis

BHC57	HK3	*Snodgrassella*	Snod_bombus	*Snodgrassella sp.*	Bombus pensylvanicus
BGI32	App2-2	*Snodgrassella*	Snod_bombus	*Snodgrassella sp.*	Bombus appositus
BGI34	App6-4	*Snodgrassella*	Snod_bombus	*Snodgrassellla sp.*	Bombus appositus
BGI36	Fer2-2	*Snodgrassella*	Snod_bombus	*Snodgrassella communis*	Bombus fervidus
SALWKB12	wkB12	*Snodgrassella*	Snod_bombus	*Snodgrassella communis*	Bombus bimaculatus
Ga0227310	ESL0253	*Snodgrassella*	Snod_bombus	*Snodgrassella sp.*	Bombus pascuorum

693

694

695 **Supplementary Table 3.** Proportions of reads mapping to species isolated in non-*A. mellifera* hosts.

696 The proportions are computed as the number of reads mapping to the species, divided by the

697 number of reads mapping to the same phylotype for this sample.

Species	**Sample**	**Proportion of reads (%) mapping out of the reads mapping to the same phylotype for this sample**
Firm5_7	N02	0.47
	N04	0.68
	N06	0.37
	N10	0.12
	N13	3.1
Gilli_bombus	N10	0.16
	N14	0.15
Gilli_apis_dorsa	F11	0.84
	F16	0.92

698

699 **Supplementary Table 4**. Matrix of the counts of the OGs annotated as H category significantly

700 enriched in the forager samples compared to the al OGs and all COG categories.

	"H" COG category	**All COG category**
All OGs	71	1779
OGs significantly enriched in foragers	18	63

701

702

703 ## References

704 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool.
705 J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

706 Ansorge, R., Romano, S., Sayavedra, L., Porras, M.Á.G., Kupczok, A., Tegetmeyer, H.E., Dubilier, N.,
707 Petersen, J., 2019. Functional diversity enables multiple symbiont strains to coexist in deep-
708 sea mussels. Nat. Microbiol. 4, 2487–2497. https://doi.org/10.1038/s41564-019-0572-9

709 Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H.,
710 Zhong, H., Khan, M.T., Zhang, J., Li, J., Xiao, L., Al-Aama, J., Zhang, D., Lee, Y.S., Kotowska, D.,
711 Colding, C., Tremaroli, V., Yin, Y., Bergman, S., Xu, X., Madsen, L., Kristiansen, K., Dahlgren, J.,
712 Wang, J., 2015. Dynamics and Stabilization of the Human Gut Microbiome during the First
713 Year of Life. Cell Host Microbe 17, 690–703. https://doi.org/10.1016/j.chom.2015.04.004

714 Barboza, M., Sela, D.A., Pirim, C., LoCascio, R.G., Freeman, S.L., German, J.B., Mills, D.A., Lebrilla,
715 C.B., 2009. Glycoprofiling Bifidobacterial Consumption of Galacto-Oligosaccharides by Mass
716 Spectrometry Reveals Strain-Specific, Preferential Consumption of Glycans. Appl. Environ.
717 Microbiol. 75, 7319–7325. https://doi.org/10.1128/AEM.00842-09

718 Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L.,
719 Marshall, M., Sonnhammer, E.L.L., 2002. The Pfam Protein Families Database. Nucleic Acids
720 Res. 30, 276–280.

721 Bendall, M.L., Stevens, S.L., Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., Schackwitz, W.,
722 Martin, J., Pati, A., Bushnell, B., Froula, J., Kang, D., Tringe, S.G., Bertilsson, S., Moran, M.A.,
723 Shade, A., Newton, R.J., McMahon, K.D., Malmstrom, R.R., 2016. Genome-wide selective
724 sweeps and gene-specific sweeps in natural bacterial populations. ISME J. 10, 1589–1601.
725 https://doi.org/10.1038/ismej.2015.241

726 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence
727 data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

728 Bonilla-Rosso, G., Engel, P., 2018. Functional roles and metabolic niches in the honey bee gut
729 microbiota. Curr. Opin. Microbiol., Environmental Microbiology * The New Microscopy 43,
730 69–76. https://doi.org/10.1016/j.mib.2017.12.009
731 Bremer, H., Churchward, G., 1977. An examination of the Cooper-Helmstetter theory of DNA
732 replication in bacteria and its underlying assumptions. J. Theor. Biol. 69, 645–654.
733 https://doi.org/10.1016/0022-5193(77)90373-3
734 Broad Institute, 2019. Picard toolkit.
735 Brochet, S., Quinn, A., Mars, R.A., Neuschwander, N., Sauer, U., Engel, P., 2021. Niche partitioning
736 facilitates coexistence of closely related honey bee gut bacteria. eLife 10, e68583.
737 https://doi.org/10.7554/eLife.68583
738 Brodschneider, R., Crailsheim, K., 2010. Nutrition and health in honey bees. Apidologie 41, 278–294.
739 https://doi.org/10.1051/apido/2010012
740 Bron, P.A., van Baarlen, P., Kleerebezem, M., 2012. Emerging molecular insights into the interaction
741 between probiotics and the host intestinal mucosa. Nat. Rev. Microbiol. 10, 66–78.
742 https://doi.org/10.1038/nrmicro2690
743 Carrow, H.C., Batachari, L.E., Chu, H., 2020. Strain diversity in the microbiome: Lessons from
744 Bacteroides fragilis. PLOS Pathog. 16, e1009056.
745 https://doi.org/10.1371/journal.ppat.1009056
746 Chen, L., Wang, D., Garmaeva, S., Kurilshikov, A., Vich Vila, A., Gacesa, R., Sinha, T., Segal, E.,
747 Weersma, R.K., Wijmenga, C., Zhernakova, A., Fu, J., 2021. The long-term genetic stability
748 and individual specificity of the human gut microbiome. Cell 184, 2302-2315.e12.
749 https://doi.org/10.1016/j.cell.2021.03.024
750 Davenport, E.R., Mizrahi-Man, O., Michelini, K., Barreiro, L.B., Ober, C., Gilad, Y., 2014. Seasonal
751 Variation in Human Gut Microbiome Composition. PLOS ONE 9, e90731.
752 https://doi.org/10.1371/journal.pone.0090731

753 De Filippis, F., Pasolli, E., Tett, A., Tarallo, S., Naccarati, A., De Angelis, M., Neviani, E., Cocolin, L.,
754 Gobbetti, M., Segata, N., Ercolini, D., 2019. Distinct Genetic and Functional Traits of Human
755 Intestinal Prevotella copri Strains Are Associated with Different Habitual Diets. Cell Host
756 Microbe 25, 444-453.e3. https://doi.org/10.1016/j.chom.2019.01.004
757 Ellegaard, K.M., Brochet, S., Bonilla-Rosso, G., Emery, O., Glover, N., Hadadi, N., Jaron, K.S., van der
758 Meer, J.R., Robinson-Rechavi, M., Sentchilo, V., Tagini, F., Class 2016-17, S., Engel, P., 2019.
759 Genomic changes underlying host specialization in the bee gut symbiont Lactobacillus Firm5.
760 Mol. Ecol. 28, 2224–2237. https://doi.org/10.1111/mec.15075
761 Ellegaard, K.M., Engel, P., 2019. Genomic diversity landscape of the honey bee gut microbiota. Nat.
762 Commun. 10, 1–13. https://doi.org/10.1038/s41467-019-08303-0
763 Ellegaard, K.M., Suenami, S., Miyazaki, R., Engel, P., 2020. Vast Differences in Strain-Level Diversity in
764 the Gut Microbiota of Two Closely Related Honey Bee Species. Curr. Biol. 30, 2520-2531.e7.
765 https://doi.org/10.1016/j.cub.2020.04.070
766 Elsayeh, W.A., Cook, C., Wright, G.A., 2022. B-Vitamins Influence the Consumption of
767 Macronutrients in Honey Bees. Front. Sustain. Food Syst. 6, 804002.
768 https://doi.org/10.3389/fsufs.2022.804002
769 Emms, D.M., Kelly, S., 2015. OrthoFinder: solving fundamental biases in whole genome comparisons
770 dramatically improves orthogroup inference accuracy. Genome Biol. 16, 157.
771 https://doi.org/10.1186/s13059-015-0721-2
772 Engel, P., Martinson, V.G., Moran, N.A., 2012. Functional diversity within the simple gut microbiota
773 of the honey bee. Proc. Natl. Acad. Sci. U. S. A. 109, 11002–11007.
774 https://doi.org/10.1073/pnas.1202970109
775 Engel, P., Moran, N.A., 2013. Functional and evolutionary insights into the simple yet specific gut
776 microbiota of the honey bee from metagenomic analysis. Gut Microbes 4, 60–65.
777 https://doi.org/10.4161/gmic.22517

778 Engel, P., Stepanauskas, R., Moran, N.A., 2014. Hidden Diversity in Honey Bee Gut Symbionts
779 Detected by Single-Cell Genomics. PLOS Genet. 10, e1004596.
780 https://doi.org/10.1371/journal.pgen.1004596
781 Garrison, E., Kronenberg, Z.N., Dawson, E.T., Pedersen, B.S., Prins, P., 2021. Vcflib and tools for
782 processing the VCF variant call format. https://doi.org/10.1101/2021.05.21.445151
783 Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing.
784 ArXiv12073907 Q-Bio.
785 Garsin, D.A., 2010. Ethanolamine Utilization in Bacterial Pathogens: Roles and Regulation. Nat. Rev.
786 Microbiol. 8, 290–295. https://doi.org/10.1038/nrmicro2334
787 Goyal, A., Bittleston, L.S., Leventhal, G.E., Lu, L., Cordero, O.X., 2022. Interactions between strains
788 govern the eco-evolutionary dynamics of microbial communities. eLife 11, e74987.
789 https://doi.org/10.7554/eLife.74987
790 Greenblum, S., Carr, R., Borenstein, E., 2015. Extensive Strain-Level Copy-Number Variation across
791 Human Gut Microbiome Species. Cell 160, 583–594.
792 https://doi.org/10.1016/j.cell.2014.12.038
793 Haiser, H.J., Gootenberg, D.B., Chatman, K., Sirasani, G., Balskus, E.P., Turnbaugh, P.J., 2013.
794 Predicting and manipulating cardiac drug inactivation by the human gut bacterium
795 Eggerthella lenta. Science 341, 295–298. https://doi.org/10.1126/science.1235872
796 Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., Bork, P., 2017.
797 Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-
798 Mapper. Mol. Biol. Evol. 34, 2115–2122. https://doi.org/10.1093/molbev/msx148
799 Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R.,
800 Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P., 2019. eggNOG 5.0: a hierarchical,
801 functionally and phylogenetically annotated orthology resource based on 5090 organisms
802 and 2502 viruses. Nucleic Acids Res. 47, D309–D314. https://doi.org/10.1093/nar/gky1085

803 Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal:
804 prokaryotic gene recognition and translation initiation site identification. BMC
805 Bioinformatics 11, 119. https://doi.org/10.1186/1471-2105-11-119
806 Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S., 2018. High throughput ANI
807 analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9, 1–8.
808 https://doi.org/10.1038/s41467-018-07641-9
809 Jones, J.C., Fruciano, C., Marchant, J., Hildebrand, F., Forslund, S., Bork, P., Engel, P., Hughes, W.O.H.,
810 2018. The gut microbiome is associated with behavioural task in honey bees. Insectes
811 Sociaux 65, 419–429. https://doi.org/10.1007/s00040-018-0624-9
812 Kaval, K.G., Garsin, D.A., 2018. Ethanolamine Utilization in Bacteria. mBio 9, e00066-18.
813 https://doi.org/10.1128/mBio.00066-18
814 Keeney, K.M., Yurist-Doutsch, S., Arrieta, M.-C., Finlay, B.B., 2014. Effects of antibiotics on human
815 microbiota and subsequent disease. Annu. Rev. Microbiol. 68, 217–235.
816 https://doi.org/10.1146/annurev-micro-091313-103456
817 Kešnerová, L., Emery, O., Troilo, M., Liberti, J., Erkosar, B., Engel, P., 2020. Gut microbiota structure
818 differs between honeybees in winter and summer. ISME J. 14, 801–814.
819 https://doi.org/10.1038/s41396-019-0568-8
820 Kešnerová, L., Mars, R.A.T., Ellegaard, K.M., Troilo, M., Sauer, U., Engel, P., 2017. Disentangling
821 metabolic functions of bacteria in the honey bee gut. PLoS Biol. 15.
822 https://doi.org/10.1371/journal.pbio.2003467
823 Kwong, W.K., Medina, L.A., Koch, H., Sing, K.-W., Soh, E.J.Y., Ascher, J.S., Jaffé, R., Moran, N.A., 2017.
824 Dynamic microbiome evolution in social bees. Sci. Adv. 3.
825 https://doi.org/10.1126/sciadv.1600513
826 Kwong, W.K., Moran, N.A., 2016. Gut microbial communities of social bees. Nat. Rev. Microbiol. 14,
827 374–384. https://doi.org/10.1038/nrmicro.2016.43

828 Leventhal, G.E., Boix, C., Kuechler, U., Enke, T.N., Sliwerska, E., Holliger, C., Cordero, O.X., 2018.
829 Strain-level diversity drives alternative community types in millimetre-scale granular
830 biofilms. Nat. Microbiol. 3, 1295–1303. https://doi.org/10.1038/s41564-018-0242-3
831 Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform.
832 Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324
833 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.,
834 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.
835 https://doi.org/10.1093/bioinformatics/btp352
836 Li, Y., Leonard, S.P., Powell, J.E., Moran, N.A., 2022. Species divergence in gut-restricted bacteria of
837 social bees. Proc. Natl. Acad. Sci. 119, e2115013119.
838 https://doi.org/10.1073/pnas.2115013119
839 Martin-Platero, A.M., Cleary, B., Kauffman, K., Preheim, S.P., McGillicuddy, D.J., Alm, E.J., Polz, M.F.,
840 2018. High resolution time series reveals cohesive but short-lived communities in coastal
841 plankton. Nat. Commun. 9, 266. https://doi.org/10.1038/s41467-017-02571-4
842 Martinson, V.G., Moy, J., Moran, N.A., 2012. Establishment of Characteristic Gut Bacteria during
843 Development of the Honeybee Worker. Appl. Environ. Microbiol. 78, 2830–2840.
844 https://doi.org/10.1128/AEM.07810-11
845 Maurice, C.F., CL Knowles, S., Ladau, J., Pollard, K.S., Fenton, A., Pedersen, A.B., Turnbaugh, P.J.,
846 2015. Marked seasonal variation in the wild mouse gut microbiota. ISME J. 9, 2423–2434.
847 https://doi.org/10.1038/ismej.2015.53
848 Moran, N.A., Ochman, H., Hammer, T.J., 2019. Evolutionary and Ecological Consequences of Gut
849 Microbial Communities. Annu. Rev. Ecol. Evol. Syst. 50, 451–475.
850 https://doi.org/10.1146/annurev-ecolsys-110617-062453
851 Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. metaSPAdes: a new versatile
852 metagenomic assembler. Genome Res. 27, 824–834. https://doi.org/10.1101/gr.213959.116

853 Patnode, M.L., Guruge, J.L., Castillo, J.J., Couture, G.A., Lombard, V., Terrapon, N., Henrissat, B.,
854 Lebrilla, C.B., Gordon, J.I., 2021. Strain-level functional variation in the human gut microbiota
855 based on bacterial binding to artificial food particles. Cell Host Microbe 29, 664-673.e5.
856 https://doi.org/10.1016/j.chom.2021.01.007
857 Powell, J.E., Leonard, S.P., Kwong, W.K., Engel, P., Moran, N.A., 2016. Genome-wide screen identifies
858 host colonization determinants in a bacterial gut symbiont. Proc. Natl. Acad. Sci. U. S. A. 113,
859 13887–13892. https://doi.org/10.1073/pnas.1610856113
860 Rao, C., Coyte, K.Z., Bainter, W., Geha, R.S., Martin, C.R., Rakoff-Nahoum, S., 2021. Multi-kingdom
861 ecological drivers of microbiota assembly in preterm infants. Nature 591, 633–638.
862 https://doi.org/10.1038/s41586-021-03241-8
863 Robinson, G.E., 2002. Genomics and Integrative Analyses of Division of Labor in Honeybee Colonies.
864 Am. Nat. 160, S160–S172. https://doi.org/10.1086/342901
865 Seeley, T., 1985. The Wisdom of the Hive: the social physiology of honey bee colonies. Princeton
866 University Press.
867 Segers, F.H., Kešnerová, L., Kosoy, M., Engel, P., 2017. Genomic changes associated with the
868 evolutionary transition of an insect gut symbiont into a blood-borne pathogen. ISME J. 11,
869 1232–1244. https://doi.org/10.1038/ismej.2016.201
870 Smith, C.R., Toth, A.L., Suarez, A.V., Robinson, G.E., 2008. Genetic and genomic analyses of the
871 division of labour in insect societies. Nat. Rev. Genet. 9, 735–748.
872 https://doi.org/10.1038/nrg2429
873 Smits, S.A., Leach, J., Sonnenburg, E.D., Gonzalez, C.G., Lichtman, J.S., Reid, G., Knight, R.,
874 Manjurano, A., Changalucha, J., Elias, J.E., Dominguez-Bello, M.G., Sonnenburg, J.L., 2017.
875 Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science
876 357, 802–806. https://doi.org/10.1126/science.aan4834
877 Sommer, F., Ståhlman, M., Ilkayeva, O., Arnemo, J.M., Kindberg, J., Josefsson, J., Newgard, C.B.,
878 Fröbert, O., Bäckhed, F., 2016. The Gut Microbiota Modulates Energy Metabolism in the

879 Hibernating Brown Bear Ursus arctos. Cell Rep. 14, 1655–1661.
880 https://doi.org/10.1016/j.celrep.2016.01.026
881 Steele, M.I., Kwong, W.K., Whiteley, M., Moran, N.A., 2017. Diversification of Type VI Secretion
882 System Toxins Reveals Ancient Antagonism among Bee Gut Microbes. mBio 8, e01630-17.
883 https://doi.org/10.1128/mBio.01630-17
884 Steele, M.I., Moran, N.A., 2021. Evolution of Interbacterial Antagonism in Bee Gut Microbiota
885 Reflects Host and Symbiont Diversification. mSystems 6, e00063-21.
886 https://doi.org/10.1128/mSystems.00063-21
887 Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., Kassabgy, M.,
888 Huang, S., Mann, A.J., Waldmann, J., Weber, M., Klindworth, A., Otto, A., Lange, J.,
889 Bernhardt, J., Reinsch, C., Hecker, M., Peplies, J., Bockelmann, F.D., Callies, U., Gerdts, G.,
890 Wichels, A., Wiltshire, K.H., Glöckner, F.O., Schweder, T., Amann, R., 2012. Substrate-
891 Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton
892 Bloom. Science 336, 608–611. https://doi.org/10.1126/science.1218344
893 Teeling, H., Fuchs, B.M., Bennke, C.M., Krüger, K., Chafee, M., Kappelmann, L., Reintjes, G.,
894 Waldmann, J., Quast, C., Glöckner, F.O., Lucas, J., Wichels, A., Gerdts, G., Wiltshire, K.H.,
895 Amann, R.I., 2016. Recurring patterns in bacterioplankton dynamics during coastal spring
896 algae blooms. eLife 5, e11888. https://doi.org/10.7554/eLife.11888
897 Tett, A., Huang, K.D., Asnicar, F., Fehlner-Peach, H., Pasolli, E., Karcher, N., Armanini, F., Manghi, P.,
898 Bonham, K., Zolfo, M., De Filippis, F., Magnabosco, C., Bonneau, R., Lusingu, J., Amuasi, J.,
899 Reinhard, K., Rattei, T., Boulund, F., Engstrand, L., Zink, A., Collado, M.C., Littman, D.R.,
900 Eibach, D., Ercolini, D., Rota-Stabelli, O., Huttenhower, C., Maixner, F., Segata, N., 2019. The
901 Prevotella copri Complex Comprises Four Distinct Clades Underrepresented in Westernized
902 Populations. Cell Host Microbe 26, 666-679.e7. https://doi.org/10.1016/j.chom.2019.08.018
903 Tsoy, O., Ravcheev, D., Mushegian, A., 2009. Comparative Genomics of Ethanolamine Utilization. J.
904 Bacteriol. 191, 7157–7164. https://doi.org/10.1128/JB.00838-09

905 Van Rossum, T., Ferretti, P., Maistrenko, O.M., Bork, P., 2020. Diversity within species: interpreting
906 strains in microbiomes. Nat. Rev. Microbiol. 18, 491–506. https://doi.org/10.1038/s41579-
907 020-0368-1
908 Whitfield, C.W., Ben-Shahar, Y., Brillet, C., Leoncini, I., Crauser, D., LeConte, Y., Rodriguez-Zas, S.,
909 Robinson, G.E., 2006. Genomic dissection of behavioral maturation in the honey bee. Proc.
910 Natl. Acad. Sci. 103, 16068–16075. https://doi.org/10.1073/pnas.0606909103
911 Wu, J., Lang, H., Mu, X., Zhang, Z., Su, Q., Hu, X., Zheng, H., 2021. Honey bee genetics shape the
912 strain-level structure of gut microbiota in social transmission. Microbiome 9, 225.
913 https://doi.org/10.1186/s40168-021-01174-y
914 Zeevi, D., Korem, T., Godneva, A., Bar, N., Kurilshikov, A., Lotan-Pompan, M., Weinberger, A., Fu, J.,
915 Wijmenga, C., Zhernakova, A., Segal, E., 2019. Structural variation in the gut microbiome
916 associates with host health. Nature 568, 43–48. https://doi.org/10.1038/s41586-019-1065-y
917 Zheng, H., Nishida, A., Kwong, W.K., Koch, H., Engel, P., Steele, M.I., Moran, N.A., 2016. Metabolism
918 of Toxic Sugars by Strains of the Bee Gut Symbiont Gilliamella apicola. mBio 7.
919 https://doi.org/10.1128/mBio.01326-16
920 Zheng, H., Perreau, J., Powell, J.E., Han, B., Zhang, Z., Kwong, W.K., Tringe, S.G., Moran, N.A., 2019.
921 Division of labor in honey bee gut microbiota for plant polysaccharide digestion. Proc. Natl.
922 Acad. Sci. 116, 25909–25916. https://doi.org/10.1073/pnas.1916224116
923 Zhu, A., Sunagawa, S., Mende, D.R., Bork, P., 2015. Inter-individual differences in the gene content of
924 human gut bacterial species. Genome Biol. 16, 82. https://doi.org/10.1186/s13059-015-
925 0646-9
926